May 2015

Informational Power on Twitter: A Mixed-methods Exploration of User Knowledge and Technological Discourse About Information Flows

Nicholas John Proferes
University of Wisconsin-Milwaukee

INFORMATIONAL POWER ON TWITTER: A MIXED-METHODS EXPLORATION OF USER KNOWLEDGE AND TECHNOLOGICAL DISCOURSE ABOUT INFORMATION FLOWS

by

Nicholas J Proferes

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Information Studies

at

The University of Wisconsin-Milwaukee

May 2015

ABSTRACT

INFORMATIONAL POWER ON TWITTER: A MIXED-METHODS EXPLORATION OF USER KNOWLEDGE AND TECHNOLOGICAL DISCOURSE ABOUT INFORMATION FLOWS

by

Nicholas J Proferes

The University of Wisconsin-Milwaukee, 2015
Under the Supervision of Professor Michael Zimmer

Following a number of recent examples where social media users have been confronted by information flows that did not match their understandings of the platforms, there is a pressing need to examine public knowledge of information flows on these systems, to map how this knowledge lines up against the extant flows of these systems, and to explore the factors that contribute to the construction of knowledge about these systems. There is an immediacy to this issue because as social media sites become further entrenched as dominant vehicles for communication, knowledge about these technologies will play an ever increasing role in users' abilities to gauge the risks for information disclosure, to understand and respond to global information flows, to make meaningful decisions about use and participation, and to be a part of conversations around how information flows in these spaces should be governed. Ultimately, knowledge about how information flows through these platforms helps shape users' informational power.

This dissertation responds to such a need by investigating the extant state of information flows on the popular social media platform "Twitter," user knowledge about information flows on Twitter, and explores how Twitter, Inc.'s messaging to users may impact users' knowledge construction. Through a mixed-method approach that includes a science and technology studies informed technical analysis of the Twitter platform, a

quantitative analysis of survey data gathered from Twitter users and non-users which tested knowledge of different aspects of information flows on Twitter, and a critical discourse analysis of Twitter's messaging to users in the new-user orientation process, this dissertation theorizes how junctures and disjunctures among the three can impact individual power. Findings of this project suggest that while many of the protocols and algorithmic functions associated with real-time information production and consumption on Twitter are well understood by users and are clearly articulated by Twitter, Inc., other aspects of information flows on the platform—such as the commodification of user-generated content, the long-term lifecycle of Tweets (such as the archival of Twitter by the Library of Congress), and the differential global flows of information—are not as well understood by users, nor explained in as much detail by Twitter, Inc. This dissertation describes the resulting state of users' informational power as one of "information flow solipsism."

Table of Contents

List of Figures

List of Tables

Acknowledgements

There are a number of people who have graciously and generously shared their time, energy, and good humor with me during this winding process. I would like to first acknowledge my chair, Dr. Michael Zimmer, for seeing me through this process and whose patience, trust, and encouragement helped me through many a difficult moment; Dr. Sandra Braman, who has continuously challenged me to broaden my intellectual horizons; Dr. Richard Grusin, who has pushed me to develop a far greater interdisciplinary perspective; Dr. Nadine Kozak, whose approach furthered my understanding of what it means to be an academic; and Dr. Wilhelm Peekhaus, whose feedback, advice, and suggestions always made me pause say, "Oh yeah... that's a really good point."

I would further like to acknowledge a number of awesome individuals who have provided valuable feedback and insights when I have been at critical junctures and who planted seedlings for ideas that have made their way into this project. This includes (but is by no means limited to): Shannon Crawford Barniskis, Ann Graf, Dr. Anna Lauren Hoffmann, Dr. John Jordan, Dr. Joyce Latham, Dr. Jeremy Mauger, and Adriana McCleer. I would also like to acknowledge my family, who has continuously encouraged me to explore while providing a great deal of sage advice and support along the way.

Finally, I would like to acknowledge Katherine Morrissey, whose love, support, and insight I cannot adequately capture on this page, but who is my beacon.

Chapter 1: Introduction

Informational Power in a Social Media Landscape

The wide array of social media platforms in existence today provides users opportunities for communication at a scale and speed that may have seemed unfathomable fifty years ago. While millions of users take advantage of these opportunities, many do not fully understand how the information they create flows through these vast, complex, and frequently opaque digital environments. For example, Acquisti and Gross (2006) found that almost a quarter of sampled Facebook users misunderstood who could access their information and Park (2013) found "more than 40% of the respondents misunderstood the most basic aspects of institutional data practices" (p. 224). Misunderstanding how social media platforms make user-generated information available to other users, how they share or sell it, or how they archive or store information can lead to serious consequences for users. Without knowledge of how information flows on these platforms, users may be limited in their abilities to understand how social media platforms filter information; to gauge the risks for information disclosure and consumption; to make fully informed decisions about use; to understand and have a say in the larger, often global information ecosystems frequently surrounding the platforms; or to participate in conversations around how information flows in these spaces should be governed. Knowledge of how information flows on social media helps make it possible for users to enter these fields of action. It functions as a form of informational power.

Braman (2006) identifies four forms of power: instrumental power, structural power, symbolic power, and informational power. Instrumental power is power that

"shapes human behaviors by manipulating the material world via physical force," structural power "shapes human behaviors by manipulating the world via rules and institutions," symbolic power "shapes human behaviors by manipulating the material, social, and symbolic world via ideas, words, and images," and informational power "shapes human behaviors by manipulating the informational bases of instrumental, structural, and symbolic power" (p. 25). In social media environments, knowledge of how information flows through a platform can function as a form of informational power. This knowledge helps individuals better understand the state of the world, it helps them understand how they might exercise other forms of power within that world, and it allows them to predict some of the potential outcomes of those actions. For example, knowing that Twitter makes Tweets available publicly by default might help an individual realize that, once they send a Tweet, those in authoritative positions would be capable of viewing it unless the individual changes the default settings. Practically then, the individual might decide based on that knowledge that making disparaging comments about bosses or superiors on the platform could be unwise. As a result of this, the individual may choose to tailor Tweets carefully or might choose to change the default settings. Conversely, the individual might gauge the risks, decide they are entirely comfortable with the possibility of discovery, and proceed regardless. In each of these cases, the knowledge of how information flows through this particular social media platform opens the possibility for different actions and informs the exercise of symbolic power on the platform. When applied, the individual's base of informational power facilitates a reduction in uncertainty and helps inform the choices they make in relation to these spaces.

In her 2012 book, *Configuring the Networked Self*: *Law, Code and the Play of Everyday Practice*, Cohen argues that, "the emerging regime of information rights and privileges … allows individuals less and less control over information flows to, from, and about themselves" (p. 3). Building on Cohen's analysis of the legal realm, I argue that having incorrect or incomplete knowledge of information flows also contributes to this end state of diminished control. Without knowing how one's data might be used, shared, stored, or archived, it becomes much more difficult for an individual to judge the potential consequences of engaging in information production or consumption processes, and much more difficult for any individual to try to exert any additional measure of power, control, or influence over information flows once the processes have been engaged. Instead of the realization of this diminished state of control occurring through law, I argue that it is also realized at the level of the individual's informational power.

Different parts of the external world shape and influence individuals' informational power. Individuals are not born into this world with wired knowledge of social media platforms. Instead, according to Rogers (2003), people build knowledge of technology over time through direct perceptions and experiences with technology, by watching others use it, and by consuming messages and communications about technology. Gill (2000) argues "linguistic forms can have dramatic effects upon how an event or phenomenon is understood" (p. 174). The information flows within a social media platform are no exception to this. The communications and messages created by a technology's purveyors can play a particularly influential role in the development of an individual's knowledge of that technology (Pfaffenberger, 1992). For example, a 2013 Pew Internet and American Life Project survey found U.S. teens report that when they

are trying to learn about the privacy settings on a social networking site (settings which allow users control of certain information flows), most of the learning that they accomplish is through a combination of trial and error with the site settings and through the knowledge gleaned from "pop-up messages and tutorials" (Madden, Lenhart, Duggan, Cortesi, & Gasser, 2013, p. 7). The potential implications for users' knowledge and the subsequent actions users might take based on the knowledge derived from this instructional language are immediately apparent. Pop-up messages, tutorials, instructional messages on websites, advertisements, and other forms of messaging generated by a technology's purveyors may each contain depictions or descriptions of the ways that information flows within the social media platform. Therefore, these texts may each contribute to and have consequences for an individual's informational power regarding that technology as they influence the subsequent decisions an individual would make on the basis of such knowledge. This discourse therefore, can help or hinder the development of an individual's informational power. While one might hope that the organizations that produce such messaging do so in a way that fully articulates the information flows of the technology, this is not always the case.

There is an inherent potential for this type of messaging to be problematic for the development of users' informational power. This is a result of the fact that this language is generated by organizations whose interests and motivations may not necessarily align well with prioritizing the development of users' informational power, particularly when the results of such informational power may include the possibility of the individual abandoning the technology or using it in a way that challenges the business owner's ability to generate profit. This raises the potential for technology purveyors to provide

incomplete, vague, inaccurate, or otherwise less than full disclosures of how information flows in these spaces in order to recruit or retain users, and to shape and structure user knowledge in the hopes of influencing their use behavior.

Morozov (2012) highlights this exact tension in his book, *Net Delusion*. He points out that much of the discourse about "Web 2.0" technologies that comes from the tech industry does not fully (or even sometimes accurately) address what happens to the information sent through these services. Morozov argues that the idealistic descriptions of these technologies given by industry leaders and technology purveyors have problematically positioned technologies, such as blogging, as inherently democratizing in order to make them more commercially successful. He criticizes this discourse because it neglects the reality that repressive regimes can use these technologies just as easily for monitoring, surveillance, and control. By not providing a full account of the potential information flows through these systems, Morozov argues users can be—and have been—put at risk. Particularly vulnerable individuals, such as political dissidents, who internalize this utopian discourse uncritically, could find themselves in dangerous positions as a result of incorrectly understanding the potential information flows.

Despite the frequent attention given to the negative impacts, results, and harms that stem from the *application* of users' misunderstandings of information flows within social media platforms, the antecedent conditions that contribute to these outcomes are not as well documented. Studies that explore how knowledge of platforms can impact users often frame the matter as an issue of digital skills rather than one of informational power. There is, however, more at stake here than just digital skills. Knowledge of what happens to the information users create on social media not only precedes and impacts

the development of digital skills, but it is also a key part of the ability to exercise power in relation to these systems. To address this gap in the body of scholarly research, this dissertation investigates the state of users' informational power on the popular social media site Twitter by surveying users' knowledge of the information flows on Twitter and analyzing how the discourse that Twitter's business owners generate depicts informational flows on the platform in relation to the actual information flows of the platform. By triangulating the three, this dissertation explores the extant state of individuals' informational power in relation to this platform.

Twitter.com

While the form of blogging Morozov identifies is still an important part of today's technological environment, Twitter's micro-blogging platform has become a significant global phenomenon. Ranked as the 8th most visited site in the world, Twitter is a unique and prolific site in the current world of Web 2.0 platforms (Alexa.com, 2015). In 2014, roughly 23% of online adults in the U.S. indicated that they have used Twitter (Duggan, Ellison, Lampe, Lenhart, & Madden, 2015). Twitter's 500 million account user base, its hyper focus on real-time communication, and its massive throughput of 140 character messages (Tweets) created and shared by users at a rate of over half a billion a day has helped make it an important site for cultural, political, and social communication.

Since the platform's founding, Twitter has become a vehicle for users to communicate, organize, respond, monitor, mediate, and even (attempt to) predict events. Political activists have used Twitter to get their message out and for coordinating activities during protests. Some of the protest events in which Twitter has been used as a communication tool include the 2007 Nigerian Election protests (Ifukor, 2010), the 2008-

2009 Iranian protests (Burns & Eltham, 2009; Grossman, 2009), and the Occupy protest movement (Juris, 2012; Thorson et al., 2013). Individuals have used Twitter to voice their pleasure or displeasure with certain products, and brand managers have used it to understand up to the moment sentiment about their products (Jansen, Zhang, Sobel, & Chowdury, 2009). Twitter has been used for detecting and tracking real-time events such as earthquakes (Earle, Bowden, & Guy, 2011; Sakaki, Okazaki, & Matsuo, 2010), natural disasters (Bakshi, 2011) and even flu propagation (Achrekar, Gandhe, Lazarus, Yu, & Liu, 2011; Lampos, De Bie, & Cristianini, 2010; Signorini, Segre, & Polgreen, 2011). Sentiment expressed in Tweets has even been used by the financial industry to attempt to predict short-term performance of the stock market (Bollen, Mao, & Zeng, 2011; Sprenger & Welpe, 2011; X. Zhang, Fuehres, & Gloor, 2011). While perhaps more mundane, many individuals use Twitter to share status updates and to engage in phatic communication (Miller, 2008). These are just a few examples from among the hundreds of uses to which Twitter has been put.

Twitter has become a prominent social media platform for online communication. However, it is also a platform where some of its users may not fully understand how information flows on the platform. For example, in 2010, the Library of Congress announced that it had struck a deal with Twitter. In a blog post entitled, "*How Tweet It Is!*," the Library declared that "Every public Tweet, ever, since Twitter's inception in March 2006, will be archived digitally at the Library of Congress" (Raymond, 2010, para. 2). With (at the time) more than 100 million users tweeting 55 million times a day (Huffington Post, 2010), Twitter had become of important cultural and historical value. Following the Library of Congress announcement, Dylan Casey, a Google product

manager commented that, "Tweets and other short-form updates create a history of commentary that can provide valuable insights into what's happened and how people have reacted" (Singel, 2010, para. 10).

Despite the potential value of a Library of Congress archive, some Twitter users were not pleased with the announcement. Comments from Twitter users on the Library of Congress' blog indicated surprise and frustration regarding the seemingly newfound permanence of Tweets. Here are three examples:

> So with no warning, every public tweet we've ever published is saved for all time? What the hell. That's awful. (Commenter-in Raymond, 2010)
>
> I can see a lot of political aspirations dashed by people pulling out old Tweets. I've always thought of the service as quite banal and narcissistic, but I've had a Twitter account to provide feedback to a college and a couple of vendors. I think I'll close my account now. I don't need to risk Tweeting something hurtful or stupid that will be around for all recorded time. (Commenter-in Raymond, 2010)
>
> Now future generations can bear witness to how utterly stupid and vain we were – 1. for creating this steaming mountain of pointless gibberings, and 2. for preserving it for posterity. LOC, you nimrods. (Commenter-in Raymond, 2011)

However, as careful observers may have already known, Tweets have never been fleeting. Twitter had always maintained in its databases all of the messages sent through its system. The company was now simply sharing this archive with the Library of Congress. However, based on these comments, it appears that some users had not perceived this to be the case. In their comments, there appears a disconnect between their perceptions of what happened to Tweets in the long-term and how Twitter was actually managing the messages sent through the service.

From this anecdote about the Library of Congress Twitter archive, let us imagine a hypothetical user. This user is operating under a false perception that Tweets are ephemeral, a perception that has been built in part on vague descriptions about the

longevity of Tweets generated by the platforms' vendors. Based on incorrect perceptions about the way that information flows on Twitter, this user may have sent a message through the system that they did not expect to be archived, let alone archived in the Library of Congress. Perhaps this message contained embarrassing, personal, or otherwise sensitive information. Perhaps this message was not even particularly noteworthy or embarrassing at the time, but in a context five years from now, becomes relevant at a job-interview. This imagined user who operated under a false perception of Twitter's Tweet storage practices, perhaps, would not have posted this message on Twitter had they understood more accurately how information flows on Twitter; if their base of informational power was more robust. Because this user did not have an accurate understanding of how information flows on Twitter, they were unable to make a more fully informed decision about use and participation. Benkler (2007) suggests, "A fundamental requirement of self-direction is the capacity to perceive the state of the world, to conceive of available options for action, to connect actions to consequences, to evaluate alternative outcomes, and to decide upon and pursue an action accordingly" (p.147). If the ability to self-direct is predicated upon our perception of the world, and if the discourse about Twitter helps shape this perception, and if this discourse was misleading, ambiguous, or unclear, individuals might be limited in their ability to set appropriate ends for themselves, vis-à-vis their diminished states of informational power.

While this anecdote and subsequent thought experiment are useful for drawing a rough sketch of the kinds of problems that are at stake, this method of analysis has shortcomings. In order to draw a more holistic picture of where there may be problems of users' knowledge of information flows on Twitter, this project must move well beyond

three comments left on the Library of Congress's announcement. In doing so, this dissertation provides a higher-resolution picture of Twitter users' perceptions of the information flows on the platform and an account of how the messages created by the business entity that runs Twitter account for information flows on the platform, with an eye towards how these depictions may potentially influence users' informational power. By exploring users' perceptions of information flows on Twitter, Twitter's descriptive language regarding the information flows of the platform, and the extant flows of the platform, this project identifies impediments users face in developing informational power in relation to this important cultural, political, and social space.

Dissertation

In this dissertation, I ask the following research question:

> In the user-Twitter relationship, what is the state of Twitter users' informational power in regards to the informational flows of the platform? What knowledge of information flows do users have and how does the technological discourse surrounding the site created by Twitter's business purveyors describe information flows and potentially impact users' knowledge (and hence informational power)?

Through a mixed-methods approach that includes both an exploratory quantitative user-study and a critical discourse analysis, this dissertation establishes an exploratory account of how user perceptions of the information flows on Twitter do or do not match up with the technical reality of information flows on Twitter, how the discourse that surrounds Twitter does or does not match up with the technical reality of information flows on Twitter, and the potential implications these junctures and disjunctures carry for user informational power on Twitter. In exploring the state of informational power among current Twitter users and in identifying the ways in which the discourse that surrounds Twitter may be shaping it, this dissertation's findings contribute to a number of different

conversations. First, they will help illustrate how incorrect or incomplete knowledge of information flows can contribute to the same state of diminished control that Cohen (2012) identifies. Second, they underscore how the discourse created by technology purveyors contains particular projections of information flows, which can have serious potential consequences for users. Finally, the findings from this dissertation will help illustrate how impediments to the development of knowledge about information flows impacts individual's abilities to exercise power in relation to these sociotechnical systems.

Structure of the Dissertation

Having introduced the research question, this section now summarizes the general structure of the dissertation. The second chapter theorizes the user-technology relationship and how power operates inside of it. Later chapters apply this theoretical lens to in order to explore Twitter users' power in relation to the sociotechnical system of Twitter. Through a synthesis of relevant literature, this chapter argues that the relationship between a user and a technology develops as the technology diffuses throughout society, and argues that the relationship inherently involves the negotiation of power in multiple forms. A particularly important form of negotiated power is what Braman (2006) identifies as informational power. This type of power serves as an informational base for other forms of power, such as instrumental, structural or symbolic power, and can be comprised by the knowledge of how a given technology functions (what Rogers [2003] refers to as *principles-knowledge* of a technology). However, factors external to the individual, such as the design of a material technology and the technological discourse that surrounds the artifact, can influence the development of an

individual's *principles-knowledge* of a technology (and hence informational power) as the individual uses these elements to build internal pictures of how technologies work in the world.

The third chapter delves deeper into a popular contemporary genre of technology: social media sites. This chapter begins by surveying how scholars have conceptualized the operation of power within the user-SMS relationship. After reviewing scholarly work that highlights both positive and negative outcomes for user power in this relationship, this chapter argues that, in looking at the negative outcomes for user power, one of the problems scholars consistently identify fundamentally relates to users' (lack of) power in relation to the information flows on these platforms. However, the body of literature that identifies this issue often does not frame it as a problem of informational power. Further, the work that highlights users' (lack of) power in relation to the information flows is often piecemeal, rarely addressing power, user knowledge of information flows, technological discourse that describe information flows, and the extant nature of the flows in tandem. From this, this third chapter argues that further study into the interconnection between these three elements is needed and posits that the social media platform Twitter as a timely space in which to investigate the interconnections.

The fourth chapter establishes a baseline technical account of how information flows across Twitter. To render visible the information flows of the platform in a methodical manner that takes into consideration both the technical elements of the system (such as data structures, algorithms, protocols, etc.) and the social elements of the system that impact information flows (such as governing documents, business practices, etc.), this chapter turns to the analytical framework put forth by van Dijck (2013) in her critical

history of social media, *The Culture of Connectivity*. Van Dijck's conceptual framework facilitates identifying and deconstructing some of the salient, yet interconnected components of social media platforms such as data and metadata structures, algorithms, protocols, defaults, informational content, users, business models, platform ownership, and governing practices. Through a close reading of the Twitter APIs and supporting documentation for application programmers, the Twitter for business webpages, Twitter's web interface, Twitter's SEC filings, the Twitter blog, the Twitter developer forms, Twitter's policy documents, and other secondary sources, this chapter articulates how the social and the technical constitute information flows on the system.

Chapter 5 operationalizes the research question of the dissertation and describes the methods used to address it. To explore the question of user knowledge, this project builds from the quantitative methodological approaches found in Hargittai (2005) and Fuchs (2009), relying on a descriptive and inferential statistical analysis of the results of a user survey that assesses Twitter users' knowledge of the data/metadata, algorithms, protocols, defaults, informational content, users, business models, platform ownership, and governing practices that constitute information flows on the platform. Subsequently, this project leverages a discourse analysis to explore how Twitter, Inc.'s language describes and positions information flows on the platform as part of the new-user orientation process.

Chapter 6 reports on the results of the user knowledge survey which was distributed to over 15,000 individuals at a large public, urban Midwestern university. Chapter 7 reports the results of the discourse analysis, which explores how the descriptive language given to users as they would experience it signing up for the platform addresses

and positions the information flows. This includes an analysis of the registration process, the terms of service and privacy policies, as well as the descriptive language of the Twitter web-interface. Chapter 8 concludes the dissertation with a discussion of the intersections and disconnects between users' understandings of the information flows on Twitter, the discourse generated by Twitter, Inc. that describes these flows, and the extant nature of information flows on the platform, with an eye towards what these intersections and disconnects may mean for users' informational power.

Chapter 2: Conceptualizing the User-Technology Relationship

Introduction

The goal of this chapter is to theorize the user-technology relationship and how power operates inside of it. Later chapters apply this theoretical lens to in order to explore Twitter users' power in relation to the sociotechnical system of Twitter. However, in order to talk about the user-technology relationship and the functioning of power inside of it, it is necessary to first discuss the context in which this relationship forms; how users and technology come together.

This chapter first argues that the process of technological diffusion can be conceptualized as the context in which users first gain exposure to technology and thereby enter into a relationship with it. During diffusion, users are exposed to a technology by consuming discourse about a technology (what this dissertation will refer to as technological discourse), watching others use the technology, or by directly encountering the technology. During this process of exposure, users build up different kinds of knowledge about the given technology, what Rogers (1962/2003) calls *awareness-knowledge*, *how-to knowledge*, and *principles-knowledge*. This knowledge about technology can influence and shape attitudes towards the technology, the use (or non-use) of the technology, and the range of affordances a user might realize through the technology. As this chapter discusses, the user-technology relationship is constituted and continuously reformed as users are exposed to a technology and a given technological discourse, and as the user constructs their knowledge in relation to these exposures.

Next, this chapter argues that the relationship between a user and a technology can be conceptualized more broadly than just as an individual in direct connection with a

material artifact. When individuals are in a relationship with technology, they are often engaging vast and sometimes sprawling sociotechnical assemblages that help give rise to and situate the technology in society. Therefore, in order to clarify what this dissertation means by "technology" in the user-technology relationship, this chapter will draw from literature in the area of science and technology studies (STS) to help explicate the definition of technology. Similarly, "users" of technology are not a singular, uniform population. There are often many different types of users who maintain relationships with technology as it diffuses throughout society and numerous ways of conceptualizing these users. As different types of users may have distinctive extant power-relations as part of their relationships with technology, this chapter also explicates what is meant by "users" as part of the user-technology relationship.

It is only once this chapter has established this larger picture of how the user-technology relationship develops and what constitutes "technology" and "user" that the chapter offers a theoretical framework for how power functions in the relationship. Drawing on definitions and conceptualizations of power offered by Lukes (1974), Foucault (1972, 1973, 1977, 1978, 1980), and Braman (2006), the final section of this chapter argues that, not only is power an intrinsic part of the user-technology relationship, but that this power is relational, that there are multiple forms of power in play in this relationship, and that these different forms of relational power are often a site of conflict. Using the concept of "informational power" offered by Braman (2006), this section will conclude by arguing that an individual's technical knowledge of a material artifact (what Rogers refers to as *principles-knowledge*) gathered through use or discourse can impact the individual's ability to exercise different forms of power in the

relationship with the sociotechnical system that artifact is part of. As a result, the user's informational power can be an important site of power struggle in the user-technology relationship.

Diffusion

In his book, *Diffusion of Innovations,* Rogers (2003) presents a theory of how innovations spread throughout society. Developed through an inductive analysis of hundreds of empirical studies on innovation adoption, Rogers argues that diffusion is "the process in which an innovation is communicated through certain channels over time among the members of a social system" (p. 5). He observes that diffusion of an innovation is, "a kind of social change, defined as the process by which alteration occurs in the structure and function of the social system" (p. 6). These innovations can be ideas, practices, or objects, which he refers to interchangeably as technology. An innovation in question need not be new to the society as a whole, but rather that, "[i]f an idea seems new to the individual, it is an innovation" (p. 12). Despite the fact that this dissertation is not explicitly interested in tracing how Twitter has diffused throughout society, but rather users' power in relation to the technology, Rogers's model is still useful framework. It helps to illustrate how the user-technology relationship develops at the individual level. Importantly, it offers a model of how individuals discover, learn about, and ultimately make decisions about use of a technology, and identifies factors within the social context in which the innovation is diffusing that influence this process.

Rogers argues that there are four factors that influence whether or not an individual adopts a given technology: the characteristics of the innovation itself, the communication channels within that society, time, and the social system the innovation is

diffusing in. He argues that first, the characteristics of the innovation, as judged by the members of the social system help to determine the adoption of the technology. Individuals may examine the properties of a new technology on the basis of that technology's relative advantages, its compatibility with the individual's existing needs, its complexity, its trialability (how much it can be piloted in individual use), and finally its observability (how well individuals can see the results of use). Rogers argues that "[i]nnovations that are perceived by individuals as having greater relative advantage, compatibility, trialability, and observability and less complexity will be adopted more rapidly than other innovations" (p. 16). Second, Rogers argues that communication channels and the messages within those channels also play a pivotal role in diffusion of a technology. These factors help propagate knowledge about a new technology and help shape attitudes towards that technology, which can influence the individual decision to adopt or to not adopt. Next, Rogers argues that time is a key variable as part of the diffusion process. Diffusion does not occur uniformly at one single moment in time, but instead is an ongoing process spread out over time. For Rogers, time is an important variable in evaluating and categorizing technology adopters. Perhaps one of the most well-known parts of his model, the adopter categories of "innovators," "early adopters," "early majority," "late majority," and "laggards" is a differentiation of users made on the basis of time technology adoption. Finally, Rogers argues that the properties of the social system the technology is diffusing into play an important role in whether or not a technology becomes widely adopted. For example, Rogers argues that social structure and social norms can dramatically influence the adoption of a technology. A technology may successfully diffuse in one setting while failing to diffuse in another as a result of

different social structures or norms. While Rogers presents each of these four factors that influence diffusion separately, he argues that they work together in conjunction in influencing an individual's technology adoption decision-making process.

Rogers defines the innovation-decision process as a five-stage progression in which, "an individual … passes from gaining initial knowledge of an innovation, to forming an attitude toward the innovation, to making a decision to adopt or reject, to implementation of the new idea, and confirmation of this decision" (p. 168). The knowledge stage is the first stage, during which the potential user first learns of a given technology and gains an understanding of how it functions. Next, the individual enters the persuasion phase, where they develop an opinion or attitude, favorable or unfavorable, about the technology based on the knowledge gained in the first stage. Next, they enter the decision phase, during which the individual decides to adopt or not adopt the given technology. If they do decide to adopt the technology, the implementation phase follows, during which the individual starts using the given technology and begins putting the knowledge gained in the first stage to use. Finally, the individual enters the confirmation stage, during which the user re-evaluates their continued use of the technology. Re-evaluation can occur as a result of new knowledge about the technology or because of a new innovation that threatens to displace the older technology. Re-evaluation does not inherently mean discontinuance of the technology in use, but that can be one outcome.

Rogers argues that the entire five-stage decision-making process "is essentially an information-seeking and information-processing activity in which an individual is motivated to reduce uncertainty about the advantages and disadvantages of the

innovation" (p. 14). Information is a critical input for the individual as part of the entire decision making process. Rogers observes that as an individual gains initial exposure to a technology, seeks out further information about the new technology, and then processes that information, that there are many different *kinds* of knowledge that individual may develop in relation to the technology. Rogers identifies three types of knowledge about technology relevant to individuals. The first of these, he refers to as *awareness-knowledge*, which is simply the knowledge that a specific technology exists. The second is *how-to knowledge*, which "consists of information necessary to use an innovation" (p. 173). Rogers notes that when a new technology is particularly complex, the amount of how-to knowledge needed for the individual to successfully adopt the technology is much higher. The third type of knowledge is *principles-knowledge*, which deals with the functioning principles "underlying how the innovation works" (p. 173). Importantly, Rogers observes that it is usually possible to adopt a technology without principles-knowledge, however, "the danger of misusing a new idea is greater and discontinuance may result" (p. 173).

Our knowledge of technology, however, does not arrive *fait accompli*. Instead, as Rogers observes, individuals build their knowledge of what a technology is, how it works, and what a given technology might offer in two distinct ways. First, an individual may learn about a technology through their direct perceptions and experiences in using the material artifact or through watching others use it. Second, the individual may also come to know a technology through stories, messages, or texts that tell them something about that technology. These messages come from the communication channels as they exist within the given social system.

Rogers focuses significant attention on the second method, elaborating how messages from mass media and interpersonal communications impact and influence adoption. However, this chapter will give equal attention to both direct exposure and communicative messages, and will detail how individuals build up knowledge about technology from exposure to each. The next two subsections of this chapter expand beyond the work of Rogers to explore theories that describe how knowledge development takes place when individuals directly experience technology or are exposed to communications about a technology. These theories provide greater detail on how individuals learn about technology, learn what technology might afford them, and learn how a technology functions, which ultimately informs the user-technology relationship.

Direct experience: perception, affordances, and mental models.

When an individual first encounters a technology, one of the elements that they focus on is the possible set of things that can be done with the technology, often called that technology's set of "affordances." The term "affordance" originally comes from the perceptual psychologist J. J. Gibson (1977), who argues that the meanings of objects in an environment can be directly perceived and that these perceptions can then be mentally linked to possible actions. For example, in perceiving a large leafy tree, the individual may observe that this object creates shade on a sunny day. As a result of perceiving this affordance within the environment, the individual may take the action of sitting down under the tree to cool off (realizing this affordance in action).

Norman (1988) and Gaver (1991) are the two authors who are generally credited for taking Gibson's concept from psychology and importing it into the study of technological artifacts and technological design. Gaver (1991) observes that any given

technology provides a set of affordances that exist in relationship with that technology's users. These affordances, "are properties of the world that are compatible with and relevant for people's interactions" (p. 79); this is to say, technologies can afford us certain interactions and outcomes within the world. However, in order for the individual to realize the affordances in action, the affordance must first be perceptible.

It is only when technological affordances are perceptible to the individual that there can be a direct link between perception and action (Gaver, 1991). When the affordances of technology are not perceivable (as when they are hidden) or are perceived incorrectly by an individual, this can lead to mistakes. Norman's 1988 book, *The Psychology of Everyday Things,*[1] is a treatise on how poor design choices can hinder the perceptibility of a technology's affordances, and hence, why badly designed technology is more likely to lead to user failures. To illustrate this point, Norman gives the example of doors that do not open the way one would expect them to based on their design. These doors are not poorly designed because they do not function; they may function perfectly reasonably as a door. They are instead poorly designed because they do not make their affordances easily perceivable, which propagates user error. When technologies are not designed to make their affordances visible, this impacts an individual's ability to achieve understanding and knowledge of the technology and to be able to connect knowledge to action and use.

Once an individual has perceived a technology, but before action, they often build a conceptual model for a technology (Norman, 1988). These conceptual models are models within the individual's mind that they use to "test" how a technological object should work. When the individual adds in the context of the environment, themselves,

[1] Later retitled: *The Design of Everyday Things.*

and other objects in relationship with the technology to this internal picture, the individual arrives at what Norman (1983) calls a mental model. Mental models are a concept from psychology, generally attributed to Craik (1943). In psychology, mental models are internal representations of the world that people use to model and predict the world around them. As imported into a context with technology, Norman (1983) argues that "[i]n interacting with the environment, with others, and with the artifacts of technology, people form internal, mental models of themselves and of the things with which they are interacting. These models provide predictive and explanatory power for understanding the interaction" (p. 7). Norman and Gaver both argue that there is an important distinction between the actual affordances of a technology and the individual's perceptions of affordances, and that a coherent mapping between these two items in mental models reduces the possibility of user error.

Mental models facilitate the hypothesizing and realization of affordances in different scenarios. However, as Norman and Gaver both note, an individual's mental models need not be fully accurate with respect to how a technology works in order to be functional. For example, an individual's mental model of a car may not include the full details of how a piston inside the car's engine works; however, that individual may still have enough other knowledge about cars to build a mental model that makes it possible to drive the car. Despite the fact that incomplete or even inaccurate mental models can still facilitate use of some technologies, the limits of an individual's mental models impact the range of affordances the individual can realize with the technology. In the car example, without the knowledge of how a piston works, the individual could still "use" the car for the purposes of driving, but might be limited in their ability to successfully re-

build the car's engine. Adding in Rogers's (2003) different forms of knowledge about technology to the work of Norman and Gaver suggests that the more detailed and accurate the individual's *principles-knowledge* of a technology is, the more detailed and accurate of a mental model he or she will have. As a result of this higher resolution mental model, the individual may be able to realize a greater range of affordances involving the technology.

Individuals do not, however, simply rely on their direct perceptions of a technology in developing mental models. Norman (1988) states of mental models that, "[w]e base our models on whatever knowledge we have, real or imaginary, naive or sophisticated" (p. 38). Individuals also learn about new technologies (Rogers's *awareness-knowledge*), how to use technologies (Rogers's *how-to knowledge*), and how technologies function (Rogers's *principles-knowledge*) through communicative practices such as talking with friends or family members, reading or seeing depictions in instruction manuals or advertisements, and through reading stories and articles about technology that are made available through mass media. This dissertation will refer to these kinds of messages as *technological discourse* and describe them in greater detail in the next section.

Technological discourse.

According to Rogers, communications about the diffusing technology within formal media channels (e.g., newspaper stories or advertisements) and within informal channels (e.g., stories from friends) are how most people learn about a new technology. In each of these channels, individuals (who at this point can be conceptualized as potential users) may be exposed to different types of messages about a new technology

that can influence *awareness-knowledge*, *how-to knowledge*, and *principles-knowledge.* These types of messages are what this dissertation will refer to as *technological discourse*.

Discourse, generally, is "language use relative to social, political and cultural formations – it is language reflecting social order but also language shaping social order, and shaping individuals' interaction with society" (Jaworski & Coupland, 1999, p. 3). Through the consumption (conscious or otherwise) of this language, individuals take in information and use it to build internal pictures of the world (Potter, 1996). These pictures in turn impact, influence, or otherwise shape individual's states of knowledge about the world (Fairclough, 1999). Discourse—as bodies of statements and language use—impacts not only what individuals understand of the world, but can also impact and influence behavior and actions within the world as this knowledge is put to use in decision-making. When discourse involves descriptions, depictions, or accounts related to a specific technology or technology in general, this language qualifies as "technological discourse."

Technological discourse can be thought of as a particular set of communicative practices involving or about technology. The sources of the technological discourse can vary. Individuals may encounter it when talking with friends or family members, reading or seeing depictions of technology in instruction manuals or advertisements, or through reading stories and articles about technology that are made available through mass media. These varied types of communication are unified by the fact that their content can detail a specific technology that is being "imagined, projected, advanced, managed, coped with, or that is emergent in the world" (Bazerman, 1998, p. 385). As technologies are

introduced and developed, they engender technological discourse as individuals begin to communicate around their existence. Bazerman writes:

> The changed conditions of life made possible by the introductions of new technology create new realms of discussion as we try to figure out what these changed conditions mean, what problems they pose, and what we can accomplish within them. Technology constantly invites social, legal, personal, and economic discussions that shape how that technology becomes incorporated into new ways of life. (1998, p. 386)

This discourse is a significant part of how material artifacts "become part of our systems of goals, values, and meaning, part of our articulated interests, struggles, and activities" (Bazerman, 1998, p. 386); how technologies become situated in our society. The attachment of meaning is how "[d]iscourse makes it possible to 'see' the object in a certain way, while limiting other ways of representing it… discourse as a relatively coherent system of meaning both enables and constrains our speaking and sense making" (Tuominen, 1997, p. 352). This is to say, technological discourse has many of its effects on the cognitive level and can be thought of as both generative and restrictive. As individuals gain exposure to technological discourse oriented around a particular artifact, it creates a tacit connection between the user and the technology as the user develops *awareness-knowledge*, *how-to knowledge*, or *principles-knowledge* in relation to the technology.

Not all messages within the broad field of technological discourse are going to be weighed and evaluated equally however, and this is where the social system itself comes into play. Messages come from a number of different speakers and through a number of different channels that may carry different levels of influence for an individual. Rogers notes that mass media can have significant influence over the decision-making process. Rogers argues that mass media are particularly influential as they can, "1. Reach a large

audience rapidly. 2. Create knowledge and spread information. 3. Change weakly held attitudes." (p. 205). Other scholars have argued that those who have the ability to shape the messages within mass media have a heightened social power. For example (and certainly not limited to just these), van Dijk (1996) writes:

> …special access to, and control over the means of public discourse and communication, dominant groups or institutions may influence the structures of text and talk in such a way that, as a result, the knowledge, attitudes, norms, values and ideologies of recipients are – more or less indirectly affected in the interest of the dominant group. (p. 85).

Habermas (2006) argues that institutional actors, such as mass media and corporations' abilities to shape public discourse is a powerful force in society, and that these actors "cannot but exert power, because they select and process politically relevant content, and thus intervene both in the formation of public opinions and the distribution of power interests" (p. 419). Cukier, Ngwenyama, Bauer, and Middleton (2009) add that, "representatives of function systems (government and corporations) and special interest groups enjoy privileged access to the media and are in a position to use professional techniques that often make them stronger than civil society actors" (p. 177). The social groups—such as business and governments—that are in the business of promoting the adoption of specific technologies often have special access to these influential communication channels. As a result, their messages may play a particularly important role influencing individual's states of knowledge regarding particular technologies.

Business organizations that purvey technological goods often attempt to speak with one coherent voice about their technological product through advertising, through press releases, and through the messages generated by the business's representatives, such as CEOs and organizational leaders (Cheney, Christensen, Conrad, & Lair, 2004).

As the messages these actors generate constitute an argument about a technology, the messages function as a form of organizational rhetoric that feeds into the larger body of technological discourse. Cheney and McMillan (1990) describe organizational rhetoric as a system of communication with a common purpose, which involves the coordinated activities of two or more persons. The organization then, "emerges and functions rhetorically through the communicative practices of its members and stakeholders" (Cheney & McMillan, 1990, p. 101).

While there can be multiple goals or outcomes of organizational rhetoric, "[o]ne function of organizational rhetoric is to try to influence *topoi* or beliefs and general assumptions held by the public [emphasis original to text]" (Cheney et al., 2004, p. 89). For example, businesses can have specific arguments about a product's salient features within an advertisement and these arguments may influence the public's beliefs or knowledge about that product. Often, these pieces of organizational rhetoric are transmitted widely through the use of mass media. Stein (2002) observes, "Popular media representations … play a central role in the hegemonic production and reproduction of perspectives on new technologies in our culture" (p. 173). While Stein focuses more on the role of advertising[2] as a particular set of messages within mass media, any accounts of a technology that a business representative provides within mass media communications can similarly function as influential argument about a technology that may impact individuals' understanding of that technology. The goal of this argument may be to impact *awareness-knowledge*, *how-to knowledge*, and *principles-knowledge,* but specifically for the purposes of encouraging the adoption of that technology and to guide its use. While different pieces of organizational rhetoric may invoke a number of

[2] Which can be a form of organizational rhetoric.

different rhetorical strategies to do this, what is of importance about these messages for the purpose of this dissertation is how this organizational rhetoric and other forms of technological discourse include certain pieces of information which can be turned into *awareness-knowledge*, *how-to knowledge*, and/or *principles-knowledge*.

It is not just the language contained within a single message that is of importance for the construction of knowledge within technological discourse, however. Bazerman (1998) suggests that it is instead the alliance of similar messaging across multiple communication channels that help construe and influence technological understanding and meaning for individuals. For example, seeing an advertisement for a particular technology in a magazine and then hearing an advertisement for the same product on the radio can reinforce the message and help influence the creation of knowledge. Mass media thus can act as a substrate for this reinforcement to occur within.

In summary, technological discourse can influence an individual's *awareness-knowledge*, *how-to knowledge*, and *principles-knowledge* of a technology. This discourse may come from many different kinds of sources, but mass media is a particularly influential source. Those actors who have privileged access to generating messages in media outlets, such as the purveyors of a technology, can influence the development of an individual's *awareness-knowledge*, *how-to knowledge*, and *principles-knowledge* regarding a technology through their messaging. The messaging of these actors, when they are business organizations, can also be understood as organizational rhetoric. Further, the reinforcement of knowledge creation can occur when messaging about a technology appears in multiple media outlets.

There are, however, two important caveats that need to be made about the relationship between technological discourse and individuals as it has been described so far. First, as Tuominen (1997) eloquently explains, "Since every discourse is related to certain social practices, no discourse, due to its intersubjective character, can be isolated from the power interests of different social groups and institutions" (p. 352). As can be tacitly seen in the conversation so far, discourse (and the control over messages within it) is intimately tied to power: a concept that has not yet addressed by this chapter. For now, this chapter will forgo a conversation about discourse and power, but will return to it in the last part of this chapter. Second, technological discourse should not be construed as operating deterministically. It is important to note that although the messages about a technology are a source of influence for an individual's *awareness-knowledge*, *how-to knowledge*, and *principles-knowledge* which feed into their adoption decisions, these messages do not function in a deterministic manner. Message receivers are not passive vessels filled purely with the messages of media, but are active sense-makers in the communications process (Windahl, Signitzer, & Olson, 2008). Communicative messages must be made sense of by a receiver who can ignore them, misunderstand or misconstrue them, can reinterpret them in any number of different ways, or may simply metabolize the message as-is.

Even when a communication is received and internalized uncritically in full, it does not mean that it will be deterministically used as part of decision-making. For example, some social cognition research has found when people construct judgments, "they typically do not search memory for all information that is relevant to the judgments, but instead retrieve only a small subset of information available" (Shrum,

2002, p. 71). Instead, Shrum (2002) notes that there are two principles that impact judgment. The first is referred to as the "Heuristic Principle" in which individuals frequently rely on what will "sufficiently" allow them to make a judgment. Sufficiency is impacted by the motivation for retrieval and the individual's capabilities for information processing. The second principle is the "Accessibility Principle." This line of social cognition theory argues that information that is most easily mentally accessible is the most likely to be used. The frequency and recency of information use, the vividness of the information, and the information's relations to other mental constructs can impact the mental accessibility of information. This underscores Bazerman's observation that repeat exposure to messaging can impact understandings of technology. For the purposes of considering users' knowledge of technology as augmented by technological discourse then, it appears that while any piece of communication that has been metabolized by the individual can contribute to *awareness-knowledge*, *how-to knowledge*, and *principles-knowledge*, the conversion of this information into decision-making and use will vary by individuals' heuristics, by the accessibility of the information, and by the individual's capacity to process information. In summary, while exposure to technological discourse can influence an individual's *awareness-knowledge*, *how-to knowledge*, *principles-knowledge*, this relationship is not one of the individual being "determined."

Moving beyond diffusion.

This review of Rogers's theory of innovation-diffusion illustrates a conceptualization of the broader context in which the user-technology relationship develops and is maintained. Rogers's work suggests that this relationship begins when individuals are either exposed to an innovation either directly or are exposed to

technological discourse about the innovation. In both of these situations, a user may develop *awareness-knowledge*, *how-to knowledge*, and/or *principles-knowledge* regarding a technology. This knowledge can affect the individual's decisions about whether or not to adopt and use a technology, and if they do adopt it, may impact the range of affordances that they are able to realize through the technology and may impact the individual's continued use of the technology. However, this knowledge development does not occur deterministically from technological discourse, nor does it deterministically direct adoption, use, or disuse of a technology.

Rogers's model of innovation-diffusion is not the only model of how technologies propagate throughout society, though it is perhaps one of the most widely cited. It has also been subject to some criticisms.[3] Of these, there are some critiques that are of particular relevance to the work of this dissertation. First, some scholars have argued that diffusion models "black-box" particular facets of the diffusion process, ignoring their sometimes messy and complicated nature, or treating parts of the diffusion process as overly linear. Although they do not identify Rogers's model by name, Pinch and Bijker (1984) make this exact critique of diffusion models broadly. Their model of technological development, the social-construction of technology model (Pinch & Bijker, 1984; Bijker, 1995), suggests that the development and diffusion of technologies do not follow inherently linear paths, but are instead much more dynamic processes, often involving multiple "relevant social groups" that shape and tweak the technology before it reaches stabilization and diffuses broadly. Further, these authors suggest that a technology should not be treated as a static entity with clearly demarcated boundaries when the picture may be much more complicated. Other scholars, such as Cowan (1987), have also suggested

[3] Rogers has responded to many of these criticisms in each updated edition of the book.

that treating technology consumers as one coherent homogenous mass is a conceptual error and that users should be studied more granularly. Thus far, this chapter is guilty of both of these fallacies, treating "technology" and "users" as static and uniform entities in order to begin the conversation about the user-technology relationship by introducing a general context in which these two come together. With Rogers's model of diffusion now introduced as the context in which the user-technology relationship develops, this chapter now delves more deeply into the constitution of "technology" and the "user." The next section of this chapter will explicate these terms, borrowing heavily from work in the area of science and technology studies.

Explicating Technology and Users

Technology as sociotechnical systems.

Orlikowski (1992) observes, "despite years of investigative effort there is little agreement about the definition and measurement of technology" (p. 398). Many use the term simply to refer to the material artifacts that humans produce, or what some call "hardware." Rogers (2003) defines technology not just as hardware but also "software, consisting of the knowledge base of the tool" (p. 36). Scholars from the field of STS, such as Hughes (1987), Bijker (1995), and Latour (2007), offer an even more greatly expanded view of what constitutes technology. Some of these expanded views include conceptualizing technology not just as hardware, but also the knowledge or skills necessary to produce the artifacts, the vast assemblages of development processes and infrastructures that help give rise to the technologies, and/or the networks of laws, economic systems and social practices that have helped to situate the hardware within society or within an organization.

In an article directed towards other STS scholars, Kline (1985/2003) argues that the discipline cannot proceed in its work without critically unpacking and examining the various meanings of "technology." Without unpacking this term, Kline argues that STS scholars will not be able to understand how innovation occurs, how culture is connected to technology, and "the way in which we humans make our living on the planet" (p. 210). This dissertation faces a similar challenge. To more fully understand the Twitter-user relationship, it needs to adopt a view of technology that can account not just for the material functioning of an artifact, but that can also help account for elements that have given rise to the shape of Twitter, such as the technology's founders, designers, and purveyors; Twitter's business practices; Twitter's third-party partners and developers; the economic environment; related technologies and infrastructure, and so on. In short, it needs to account for the things that have helped to produce and situate the artifact in society today. Because these contextual elements contribute to the shape of the material artifact, they also inherently impact the subsequent relationship between the user and the material artifact.

Kline suggests that there are essentially four distinct ways that technology can be understood: as an artifact, as sociotechnical systems of manufacture, as technique, and as sociotechnical systems of use. In the artifact view, technology is defined simply as material objects such as manufactured products, devices, systems, equipment, machines, instruments, or other forms of worked objects or physical hardware. For example, the material object of a car would be considered as an artifact. This material artifact, of course, is still a complex object and can be further broken down into smaller artifacts, such a tires, engine, pistons, frame, seats, etc. Kline, however, argues that this view of

technology is overly reductive. For example, it ignores the technical knowledge, skills, and related infrastructure necessary to produce a technological artifact. It also ignores what Feenberg (1992) calls "technical codes," the social mediation that often governs the construction of artifacts, such as laws that might require seatbelts to be included in cars. These entities—technical knowledge, skills, infrastructure, and technical codes—are also human made artifacts, albeit not necessarily physical ones, and are often either a necessary precondition for a material artifact to come into existence or can drastically shape the technology's final form.

Kline moves on to suggest a second view of technology that accounts for the elements needed "to manufacture a particular kind of hardware, the complete working system including its inputs: people; machinery; resources; processes; and legal, economic, political and physical environment" (2003, pgs. 210-211). This sociotechnical system of manufacture view considers technology as a set of systems required for producing a material artifact. In this view, all of the inputs required for the production of the car, such as the factory, assembly line, robotic equipment, engineers, drafters, lawyers, technical codes, labor agreements, and so on, would be considered as part of the "technology." However, even in this view, there are still significant elements not accounted for.

The third view expands beyond the physical object and the sociotechnical system of manufacture to consider the technical knowledge and skills required to produce the material artifact (Gehlen, 2003). Kline (2003) traces this definition in part back to the work of Ellul (1964) who uses the term technology to refer to any form of "rationalized methodology" (a phrasing that builds on Max Weber's analysis of technology). However,

the Greek root of the word technology, *techné*, also captures this range of referents. *Techné* specifically includes knowledge and action related to craft in addition to material artifacts (Rooney, 1997, p. 402). In the car example, this view of technology might prompt the consideration of the tacit or codified knowledge or technique of the drafter, engineer, welder, painter, and so on; sets of knowledge that contributes to the production of the car.

The fourth and final view of technology that Kline presents is that of technology as a sociotechnical system of use. The sociotechnical system of use view considers technology as the basis for "what we do with the hardware after we have manufactured it" (Kline, 2003, p. 211). It includes the elements that help situate the material artifact in society. In the car example, this view of technology might prompt consideration of society's "system of roads, gas stations, laws for ownership and operation, rules of the road, etc." (Kline, 2003, p. 211). This view begins to speak to the ways that technologies are socially produced and embedded in daily life. Kline argues that, in order to be effective, STS scholars must account for more than just hardware in order to understand the extension of human capabilities through technology. Kline argues that without accounting for the contextual factors of use, hardware is meaningless and without purpose (2003, p. 211).

Ultimately, Kline advocates for STS scholars to account for the artifact view, the sociotechnical systems of production, the technique required to produce the artifact, and the sociotechnical systems of use in their studies using the broad amalgamation of "sociotechnical systems." While not necessarily responding directly to Kline, numerous STS scholars have developed accounts of the contextual factors surrounding technology

in addition to the materiality of technology, such as in Pinch and Bijker's (1984) social construction of technology (SCOT) approach, Hughes' (1986) seamless web approach, and Latour's (2007) actor-network (ANT) approach.

For the purposes of this dissertation, embracing a sociotechnical systems view of Twitter makes more sense than rather just an artifact view. This view of technology facilitates considering how a Twitter user enters a relationship with not just as a material artifact in moments of use, but also makes room for considering how the user then exists in relation to Twitter's owners, developers, business practices, production processes, and techniques that have shaped the production of the material artifact, as well as the contextual elements that have helped situate Twitter in society since its creation. All of these factors are relevant to this dissertation's inquiry as they ultimately shape the relationship between a user and Twitter. However, as discussed next, similar to the explication of the term "technology," it is also necessary to delve deeper into what is meant by the term "user."

The multiplicity of users.

Users are not a homogenous mass that acts uniformly. They may be individuals or organizations. They may have wildly different needs, skills, demographic characteristics, social statuses, and can have varying communication practices. Among the various approaches to understanding technological innovation and diffusion discussed so far, each has approached treating and classifying the user differently, often for different ends. This section will provide a brief review of how users have been conceptualized in work that focuses on the user-technology relationship before moving to discuss how these works can inform this dissertation's understanding of users and Twitter.

Rogers's (2003) work on the diffusion of innovation treats users in a number of different ways. First, Rogers focuses on "adopters" and not "users" in his work. The term "adopter" is important in Rogers work as he sets up a binary distinction between adoption and non-adoption, as these are the end points of the individual's innovation-decision process. This binary distinction is found throughout Rogers's work stems from his focus on understanding how the consumption or non-consumption of a technology spreads across a society over a given period of time.

Eventually, this binary becomes more nuanced as Rogers introduces the variable of time, differentiating among those who choose to adopt a given technology by classifying them according to the time-sequence in which they made their decision to adopt. Based on a statistical interpretation of an observed S-shaped curve of adoption, Rogers argues that there are five adopter categories: innovators, early adopters, early majority, later majority, and laggards (p. 280). He argues that broad generalizations can be made about each category of adopters and that, notably, there are key differences in these groups in terms of their socioeconomic status, personality, and communication behavior. First, he argues that innovators are the earliest of all adopters. These adopters are obsessed with trying new innovations, are often risk takers, and often have complex technical knowledge that facilitates adoption. The next group in the time-sequence is the early adopters, who are often "opinion makers" within the social-context the innovation is diffusing in. These individuals are often generally open to the changes brought on by innovation. The early majority follows the early-adopters and are much more deliberate in their adoption choices, often interacting with other early majority members, "but seldom hold positions of opinion leadership in the system" (p. 253). They often require

more convincing regarding the merits of a particular innovation than early adopters. Late majority members are the adopters that follow after more than half of the total population of use has adopted the technology. These individuals have a tendency to be more skeptical of innovations, often choosing to adopt because of either social pressure or economic necessity. Finally, laggards are the last chronological group of adopters. Rogers argues that these latecomers are the most socially isolated of all adopters and are often suspicious of innovation.

Two critiques can be made of this view of "adopters." The first is that this view treats adopters within the five categories as fairly uniform when there may be important differences in how users act within the groups. Second, this view does not pay much attention to non-adopters, does not detail why those individuals are non-adopters, and treats non-use somewhat uniformly. In a chapter entitled, "Non-Users also matter: The construction of users and non-users of the Internet" Wyatt (2005) argues that use and non-use need to be examined more robustly when studying society and technology. Wyatt first argues that, like Rogers, one can differentiate between "current user" and "non-user." Wyatt argues that among those who do qualify as "current users" that there is a continuum that exists in terms of degrees of participation or use and that these degrees of participation or use may change over time. Accounting for different kinds of use patterns among users can help illustrate how the connection between user and technology varies in relation to technological practices. Further, among the groups who are "non-users," there may be a variety of reasons why these actors are non-users. Wyatt argues that researchers can distinguish between four types of non-users: "resisters" who have never used a given technology, "rejecters" who stopped using a technology voluntarily,

"the excluded" who do not use a given technology because they cannot gain access, and finally "the expelled" who have stopped using the given technology involuntarily either because of the cost or because they lost access for another reason. Accounting for a variety of non-users opens the door to exploring why users might leave the user-technology relationship or exploring how and why individuals never enter into the relationship. As Wyatt puts it: "[i]ncluding the variety of non-users also helps to open the way for subtler description and analysis of multiplicity of users" (p. 77). While this approach conceptualizes users and non-users differently than the work of Rogers, it still conceptualizations the relationship in terms of consumption/non-consumption.

In his book *Of Bicycles, Bakelites, and Bulbs: Toward a Theory of Sociotechnical Change*, Bijker (1995) sets out to offer a heuristic for tracing technological development and sociotechnical change. He argues that in studying how technologies become part of society, "one should never take the meaning of a technical artifact or technological system as residing in the technology itself. Instead one must study how technologies are shaped and acquire their meanings in the heterogeneity of social interactions" (p. 6). Perhaps unsurprisingly given this statement, Bijker focuses heavily on how social actors influence the shape and meanings of a given technology. However, in his work, rather than focusing strictly on users as consumers, Bijker positions them as active parts of the technological construction process. He argues that users are often a "relevant social group," a key theoretical construct for understanding the development and evolution of technology. However, while users and different subsets of users are often important relevant social groups in Bijker's work, these groups must also be placed in study among

other relevant social groups, such as the technology's inventors, advertisers, politicians, and more.

Bijker suggests that the identification of "relevant social groups" is a key stage of the research process. He proposes a two-step methodological process to identify the key players: "rolling a snowball" and "following the actors." "Rolling a snowball" (according to Bijker) is a multi-step process in which a researcher first identifies relevant actors by tracing their mention in either interviews or in reading historical documents. Next, the researcher attempts to trace how that relevant actor identifies *other* relevant actors. For example, in an interview with a technology designer, not only would the technology designer count as a "relevant social group," but the researcher would also ask the technology designer to identify other individuals or groups that are important to the technology. In historical document analysis, this is a bit more difficult and involves tracing how one group of actors discusses other actors. For example, Bijker notes that when looking at the history of the bicycle, bicycle advertisers (one important "relevant social group" identified through a search of historical documents) identify "rich, young, athletic men" as additionally relevant to the object. When a researcher does not find any references to new groups, then the initial identification process is complete.

Once such a listing of relevant groups is compiled, Bijker argues that a researcher must next "follow the actors," attempting to learn about the relevant groups in greater detail. As part of this process, the researcher should attempt to figure out how the actors delineate one group from other groups, as there is often overlap among the identified relevant social groups. During this process of boundary tracing, Bijker notes that the researcher may have to re-draw the lines around relevant groups:

> …the boundaries of social groups, although once clear-cut, may become fuzzy; new groups may split off and old groups may merge into new ones. Actors thus "simplify" and reorder their world by forgetting about obsolete distinctions or by drawing new boundaries. (p. 48)

Tracing the boundaries between these groups is often a matter of tracing how the groups differentiate each other from their own relative perspective. For example, Bijker notes that bicycle producers (a relevant social group) differentiated between bicycle non-users in general and non-users that were women specifically.

One of the benefits of this approach to identifying relevant social groups is that it treats the identified relevant social groups equally. For example, if identified as relevant, non-users might be treated with equal importance as users or designers. Further, the differentiation made among groups is not purely that of the researcher, but is also based on the perspective of the other relevant social groups. This allows one to consider, for example, how a technology's creators conceptualize and define users and non-users. One of the difficulties of this approach, however, is that it can overlook disempowered social groups, groups that are not discussed. Bijker appears aware of this critique, noting that, "those that do not have the ability to speak up and let themselves be found by the analyst-will thus be missing in the account" (p. 48); however, he does not offer a solution to this problem.

This review of approaches to users suggests that, first, it is important to recognize that there are a multitude of users and non-users, and that these individuals may have significantly different relationships with a given technology. Rogers's work suggests that time of adoption may be an important characteristic of the individual to consider when exploring a user-technology relationship, particularly as early-adopters may have more pre-existing technical knowledge and as they may be more willing to try new

technologies. Wyatt's work suggests that different use practices may be an important characteristic of the individual to consider when exploring a user-technology relationship. Further, Wyatt's work on non-use suggests that non-users should also not be treated uniformly. There are multiple reasons why an individual may not use a service, such as active rejection, resistance, exclusion, or expulsion. Finally, Bijker's work suggests that how one relevant social group conceptualizes other relevant social groups may also be an important consideration when categorizing users. For example, how a technology's purveyors define and stratify users may be an important heuristic for differentiating among users.

This dissertation will return to discussing the specific different types of users on Twitter more fully in Chapter 4. For now, this review has been provided in order to discuss how—for the purposes of establishing our theoretical lens—"users" in the user-technology relationship may refer to a wide range of referents, individuals and uses, and that "users" of technology are not a singular, uniform population. Further, in using the term "user" this work does not mean to treat "non-users" as a singular undifferentiated category. There are often many different types of individuals who maintain relationships with technology as it diffuses throughout society. These individuals may have different levels of exposure directly to a technology and to the discourse that surrounds a technology, different states of knowledge regarding technology, and different patterns of use (or non-use) regarding technology.

Drawing a picture of the user-technology relationship.

What this review of conceptualizations of "users" and "technology" demonstrates is that there is much more going on underneath the surface of these terms, that this

dissertation must take care in being specific about how "users" and "technology" are conceptualized when discussing Twitter and the individuals associated with the platform. This is both a conceptual and methodological concern. To provide a summary of the picture of the user-technology relationship that this dissertation has provided thus far, first, this relationship is one that develops during the process of diffusion. It is made extant as an individual becomes exposed to a technology directly or to technological discourse. In both cases, an individual may develop *awareness-knowledge*, *how-to knowledge*, and/or *principles-knowledge* in relation to a technology. This knowledge can contribute towards the formation of attitudes regarding the technology, may impact adoption (or non-adoption) and use decisions, and can inform the range of affordances an individual can realize in use of the technology. Further, in this relationship, a user may not be in connection with just a material artifact, but also an entire sociotechnical system. However, there are many different kinds of users that exist in relation to a sociotechnical system. A drawing of the user-technology relationship as described so far is provided in Figure 1.

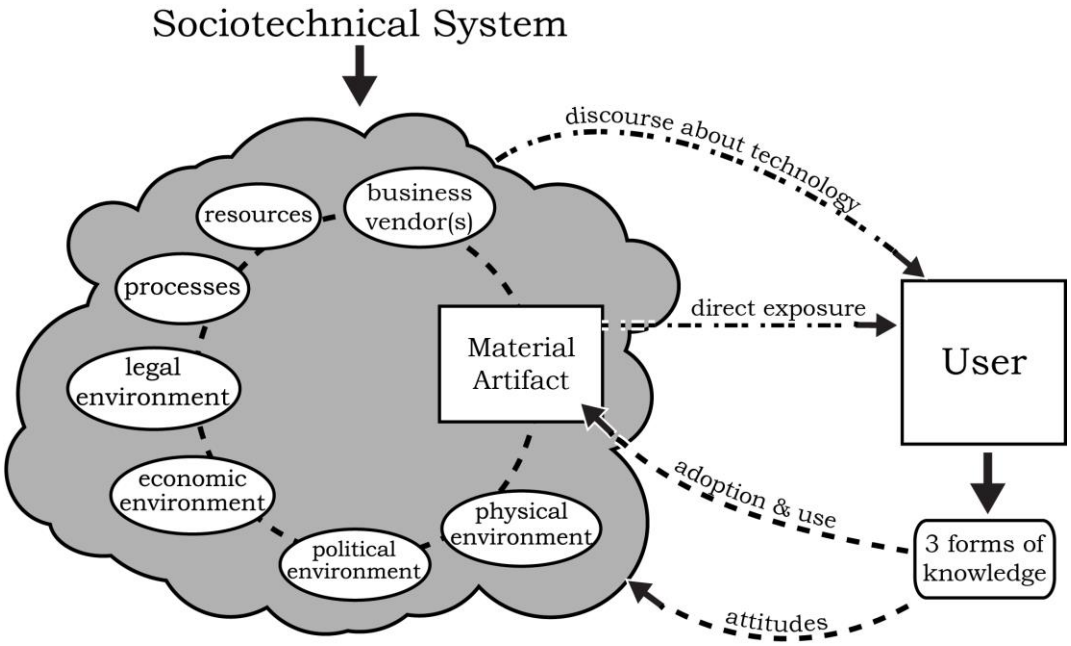


Figure 1. Model of user-technology relationship.

With this broad picture of the user-technology relationship now established, this chapter now moves to its final section to discuss how power can be conceptualized as operating within this relationship.

Power

This section argues that the relationship between an individual and a technology (as a sociotechnical system) can be conceptualized as inherently involving power. However, similar to the terms "technology" and "user," "power" is also a term that has been defined in a numerous ways by scholars in different fields of study and can refer to quite different phenomenon. The elusiveness of a single comprehensive definition led the sociologist Talcott Parsons (1963) to state that power is "a concept on which, in spite of its long history, there is, on analytical levels, a notable lack of agreement both about its specific definition, and about many features of the conceptual context in which it should be placed" (p. 232). In order to more fully articulate how the relationship between an

individual and a technology involves power, this section will review how a selected set of scholars have defined power, how they have conceptualized the different dimensions of power, how they have hypothesized how power functions within relations between actors, and how they have hypothesized what forms power may take. As the goal of this section is not to provide a comprehensive account of the history of thought on power, but rather to operationalize a view of power that can inform this dissertation's approach to understanding and studying the relationship between an individual and a technology, much of the history of thought on power presented here is abbreviated.

Defining power.

As mentioned previously, one of the fundamental problems in defining power is the difficulty of capturing all the potential forms that power may take and all the contexts in which power may exist. Scholars such as Haugaard (2002) and van Dijk (1989) warn away from reducing such a complex phenomenon to a single totalizing definition, stating respectively that there "is no single definition of power which covers all usage" (Haugaard, 2002, p. 1) and "power cannot simply be accounted for in a single definition" (van Dijk, 1989, p. 19). The hazard of adopting a single definition is that when these kinds of definitions are offered, they are often deployed to reinforce certain theoretical positions (Lukes, 1974) and thereby may leave critical elements out of consideration. If one adopts a single view of power, then one is inherently risking also adopting a particular worldview that may ignore alternative forms of power or the historical power structures that give rise to the current ones (Cameron & Kulick, 2003). To that end, Foucault warns against an overarching theory of power, arguing, "If one tries to erect a

theory of power one will always be obliged to view it as emerging at a given place and time and hence to deduce it, to reconstruct its genesis" (1980, p. 199).

Despite warnings away from grand theories of power, the term has been operationalized in a number of ways across different disciplines, including political science, communications, and sociology. Across the definitions mobilized in these areas of study, there are some consistencies as to what power enables and how power functions. For example, Max Weber, who many cite as one of the starting points for contemporary thinkers on power (Braman, 2006), is famously quoted as defining power as "imposing one's will upon the behavior of other persons" (as cited in Bendix, 1978, p. 290). Many definitions of power rally around a similar phrasing, describing power as the ability to achieve an outcome, despite another actor's wishes. For example, Parsons (1963) states that power is "the capacity of persons or collectives 'to get things done' effectively, in particular when their goals are obstructed by some kind of human resistance or opposition" (p. 232) and Haugaard (2002) offers that "power entails the capacity of one actor to make another actor do something that they would not otherwise do" (p. 304). To put this operationalization of power more formulaically, power is the ability of actor A to get actor B to do what actor A wants, despite actor B's wishes otherwise. While power is perhaps most easily understood and empirically observed within this particular arrangement, Lukes (1974) argues that there are significant functions and dimensions of power left out by this "one-dimensional view".

The dimensions of power.

Lukes observes that this "one-dimensional view" of power only seems to function in the situational context of an observable conflict where two actors—actor A and actor

B—have separate interests, and yet one actor manages to have their interests prevail within the context of interaction and conflict. Formulated this way, power only seems to help accomplish domination of one actor over another where that domination is measured *ex post facto*, through weighing outcomes against each actor's initial interests. Lukes argues that this reductive view of power ignores the ability of the more empowered actor to prevent the less empowered actor's interests from ever possibly arising within an explicit conflict in the first place. Focusing on power within the process of democratic decision-making, Lukes observes that this one-dimensional view of power ignores actor A's ability to function as an agenda setter and to exercise the power to control the context in which conflict would arise. The ability to effectively set the rules of the game must be considered in addition to the first dimension of power in order to arrive at what Lukes calls a "two-dimensional" view of power. In two-dimensional power, actor A, often through non-decision, can prevent the wishes of actor B from ever being expressed within conflict. For example, a politician's ability to determine what will and will not be talked about at a meeting is an expression of power in this second dimension. This view of power, while still outcome based, expands what actions can be considered as the exercise of power. However, as Lukes notes, power in this dimension is much more difficult (though not impossible) to empirically observe.

Lukes argues that even this two-dimensional view of power misses key functions and outcomes of power and is therefore incomplete. He writes that it does not account for the fact that actor A may be able to manipulate actor B's interests in such a way that actor B does not recognize or act on their 'real' interests. He refers to this as the capacity to produce "latent conflict, which consists in a contradiction between the interests of those

exercising power and the real interests of those they exclude… These latter may not express or even be conscious of their interests" (p. 24-25). Lukes calls this particular facet the "third-dimension" of power. Haugaard (2002) argues that Lukes' formulation of the third dimension of power can be understood as the power to produce false consciousness.[4] Rather than power resulting in domination as achieved at the site of explicit conflict or through control of the conflict environment, this third dimension of power works much more tacitly, again avoiding conflict (as in two-dimensional power) by operating at the location of actor B's cognition. It involves the manipulation of actor B's interests through the work of ideology deployed in discourse and through the manipulation of knowledge or understanding. Lukes argues that this exercise of power can occur "through the control of information, through the mass media and through the processes of socialization" (p. 23). The site of operation for power expands in this third-dimension to now include spaces such as mass media and other venues of discourse that can act as vehicles for impacting cognition. Again, Lukes observes that this dimension of power is much more difficult (but again, not impossible) to observe empirically than the first dimension.

Lukes' three dimensions of power can apply to the relationship between an individual user and a technology (as a sociotechnical system). For example, at the first dimension of power, an actor within the user-technology relationship may have power if one is able to get the other actor to do what first actor wants, despite the second actor's wishes otherwise. One context for this conflict is where a user is in direct use of the

[4] A theory originating with Marx in which the working class is essentially cognitively manipulated through ideology (or in Gramsci's [1995] work, hegemony) so that they do not recognize and act on their 'true' interests, and instead act in the interests of the capitalist class. This theory has also been criticized heavily, see for example Boudreaux & Crampton (2003) and Scott (1990).

material artifact. For example, the sociotechnical system of Twitter has a certain kind of power over users based on the fact that the shape and configuration of the material technology may limit the set of possible actions a user can take in using the technology. For instance, through code, Twitter famously limits the number of characters allowed in a Tweet to 140. A user who wishes to compose a single Tweet with 260 characters will find that they simply not empowered in their use relationship with Twitter to do this. In this way, Twitter's engineers' and designers' abilities to structure the material conditions of the technology of Twitter becomes a way of shaping and regulating a user's possible interactions with the material technology in ways that may run counter to a user's wishes. Thus, the ability to control the shape of the technology becomes a mechanism for one-dimensional power to be exercised. More broadly than the example of Twitter, scholars such as Lessig (2006) have argued that this exact control over "code" is a powerful tool in the regulation of individual behavior and use of a technology.

At the second dimension of power, a technology (as a sociotechnical system) might have power in relation to a user (or vice-versa) if one actor in this relationship can prevent the wishes of the other from being expressed within a conflict through the control of the conflict environment. In this sense, the ability to "set the agenda" becomes an important factor in the exercise of power. One example of this specific to the Twitter-user relationship is the business that runs Twitter's ability to set the agenda for the kinds of considerations or concerns that manifest into specific changes within the material technology. As a business enterprise, Twitter's leaders, engineers, lawyers, and designers are ultimately are the ones that make decisions about what design features should be implemented into the material technology or which matters are addressed in governing

documents such as Terms of Service and Privacy Policies. While user wishes and desires sometimes manifest into changes in code or policy, such as when Twitter adopted the user-generated conventions for re-Tweeting (Stone, 2009b), the technology's purveyors are the ones that ultimately control the fora through which grievances are aired, changes are made, and decisions are made about which user desires are heard and which are not.

At the third dimension of power, a technology might have power over a user (or vice-versa) if one actor is able to manipulate the other's knowledge or understanding "through the control of information, through the mass media and through the processes of socialization" (Lukes, 1974, p. 23), thereby acting on the other's cognition. As applied to the relationship between Twitter and a user, if Twitter's purveyors have the ability to control information about the material technology through the mass media and through the processes of socialization, this ability may be a vehicle for shaping the knowledge, beliefs, or attitudes that users maintain in relation to the technology. In turn, through shaping knowledge, beliefs, or attitudes towards the technology, Twitter's purveyors may be able to influence the behavior of individuals in their use of the technology.

Across all three of these dimensions, power can be conceptualized as a repressive or dominating force when successfully exercised, even though the site of the exercise of this power can vary. However, as described next in reviewing some of the work of Michael Foucault, power can also be conceptualized as a positive and generative force. Foucault's work suggests that even if and when one actor is able to achieve domination over another, it does not mean that the dominated actor is completely disempowered; both actors still maintain power in the context of the relationship.

Relational power.

Despite considerably expanding the conceptualizations of power that came before him, many have critiqued Lukes' outcome based definitions of power as it positions power as a repressive or negative force. The conceptualization of power as a restrictive force is driven in large part by both Lukes' and his predecessors'[5] desire for an empirical methodology for observing and measuring power purely within a political context. Others scholars, such as Michel Foucault, have suggested that understanding power as a purely linear and negative force misses the broader picture of power as relational and neglects what power is capable of producing.

While Foucault never offers a single comprehensive theory of power[6], across his many works (1972, 1973, 1977, 1978, 1980), he observed how particular formations or strategies of power function to not just discipline individuals, but to also produce things such as knowledge, truth, and subjectivity. Eschewing a view of power that focuses purely on repression, Foucault asked how individuals and particular states of knowledge within individuals are produced. He describes his line of inquiry this way, "… rather than ask ourselves how the sovereign appears to us in his lofty isolation, we should try to discover how it is that subjects are gradually, progressively, really and materially constituted through a multiplicity of organisms, forces, energies, materials, desires, thoughts, etc." (1980, p. 97). Foucault suggests that it is through this constitution of subjects that we can find power and its effects. In some ways, Lukes' third dimension of power seems aligned with Foucault's work. Both are interested in how power relates to the production of knowledge, thoughts, desires, and the interests of actors. However,

[5] Such as Dahl (1957), Bachrach & Baratz (1962).
[6] And Foucault rejects the notion that there should be a single comprehensive theory of power.

unlike Lukes, rather than viewing this as purely a negative and repressive force, Foucault observes that it instead productive.

Foucault's conception of power also extends it beyond the unilinear form of actor A getting actor B to do what actor A wants, despite actor B's wishes otherwise. Instead, Foucault suggests that power must be viewed as something relational. He writes:

> [P]ower is not to be taken to be a phenomenon of one individual's consolidated and homogenous domination over others, or that of one group or class over others. What, by contrast, should always be kept in mind is that power, if we do not take too distant a view of it, is not that which makes the difference between those who exclusively possess and retain it, and those who do not have it and submit to it. Power must be analysed as something which circulates, or rather as something which only functions in the forms of a chain… Power is employed and exercised through a net-like organization. And not only do individuals circulate between its threads; they are always in the position of simultaneously undergoing and exercising this power. (1980, p. 98)

In conceptualizing power as a net in which actors are situated, Foucault also inherently argues that even when actor B is 'dominated' by actor A, actor B is both undergoing and also simultaneously exercising power. This is to say, even those that are dominated are still part of power relations and still part of the network of power (Dreyfus & Rabinow, 1982). Further, in this formulation of power, Foucault argues that power only exists in relations and not as a resource that operates independently of context. He states "Power is not possessed, it acts in the very body and over the whole surface of the social field according to system of relays, modes of connection, transmission, distribution, etc." (1979, p. 59). But if this is the case, an important question follows: what are the modes of connection that sustain this net of power and how is does this work in the user-technology relationship?

Across Foucault's work, multiple 'substrates' or strategies of power are offered as ways power relations are maintained. This section will briefly focus on one such

substrate: that of discourse. Discourse, for Foucault, is not language in itself, but rather systems of representation (Hall, 2001). Discourse, as defined earlier, "is language use relative to social, political and cultural formations – it is language reflecting social order but also language shaping social order, and shaping individuals' interaction with society" (Jaworski & Coupland, 1999, p. 3). Summarizing Foucault's many works on discourse and power, Hall writes:

> Discourse, Foucault argues, constructs the topic. It defines and produces the objects of our knowledge. It governs the way that a topic can be meaningfully talked about and reasoned about. It also influences how ideas are put into practice and used to regulate the conduct of others. (2001, p. 72)

Power, as embedded throughout discourse, helps create and regulate knowledge, which in turn influences and shapes the behavior of individuals. As a result of this substrate of discourse, in Foucault's view, power and knowledge are intimately linked.[7] Power is generative for knowledge, and there is no knowledge without power. He writes, "power and knowledge directly imply one another; that there is no power relation without the correlative constitution of a field of knowledge, nor any knowledge that does not presuppose and constitute at the same time power relations" (1977, p. 27). Understood this way, language, and the use of language, does not function in a neutral way. Instead, the use of language is a key part of the exercise of power as it influences the construction of knowledge and what can be "known."

To explore the relationship between power, discourse, knowledge and the individual, Foucault engages a historical analysis[8] of how medical discourse "constructed" madness and the mad as subjects of knowledge (1988), how prisoners' subjectivities became the target of discipline (1977), and how various actors within social

[7] He goes so far as to refer to this as "power/knowledge".
[8] Referring to his method in his early work as archeology and in his later work as genealogy.

institutions such as education, psychology, police, and government used discourse to produce sexuality as an object of knowledge and to then regulate it (1978). As these discourses govern what can be considered factual or a "true" statement in a given context, this language use has a normative function in relation to knowledge. Foucault argues that, once internalized, discourse has normative effects on the individual, inherently shaping how an individual subjectively sees and experiences the world. This subjective knowledge then becomes the basis of social practices that then further inform new power relations.[9] As a result, discourse functions as a critical substrate and method through which power is practiced, sustained, and reproduced. However, even in its discursive form, power is not absolute for Foucault. He states in his book *The History of Sexuality, Vol. I*, discourse "transmits and produces power; it reinforces it, but also undermines and exposes it, renders it fragile and makes it possible to thwart it" (1978, p. 100).

Foucault's observations about power can be applied to this dissertation's conceptualization how power functions within the user-technology relationship. For example, in light of Foucault's arguments, one might ask how a technology's purveyors exercise power when creating technological discourse, how this discourse/power helps produce certain states of knowledge within individuals, and ultimately, how it helps produce "users." Foucault's work suggests that technological discourse can draw individuals and sociotechnical systems together even in the absence of an individual's use of the material artifact as it influences the individual's knowledge of the technical artifact. It suggests that these prescribed states of knowledge may become productive for users as they apply this knowledge to artifact adoption, use, or towards other ends.

[9] And hence, here, we can see why Foucault eschews a totalizing definition of power, because "If one tries to erect a theory of power one will always be obliged to view it as emerging at a given place and time and hence to deduce it, to reconstruct its genesis" (1980, p. 199).

Specific to the Twitter-user relationship then, one might ask how the discourse generated by Twitter constitutes the material artifact as an object of knowledge, how this discourse is internalized by the individual user and turned into operational knowledge, how this operational knowledge then informs a user's power, and how this user's power then informs the user's interactions and use of the technology.

Foucault's work also suggests that even if one actor in the user-technology relationship is able to achieve dominance over another within a particular conflict, this power should not be construed as absolute. Instead, each actor maintains relational power that is productive. Further, Foucault's observations about resistance suggest that, in the user-technology relationship, while technological discourse created by Twitter's purveyors may influence and shape knowledge, this discourse can be manipulated, or perhaps even undermined.

The forms of power.

While Foucault's observations about the positive and relational aspects of power are important for expanding our conceptualization of the relationship between a user and a technology, the forms of power he identifies (such as disciplinary power) often vary in relation to the context of his study.[10] In conceptualizing the relationship between a technology and a user as a relationship where each respective party maintains power in relation to the other, it is therefore important to ask: what forms of power could one expect to observe in this relationship? To answer this question, this chapter next turns to a different scholar who provides a taxonomic approach to accounting for the forms of power.

[10] And not without reason, after all, different power relations are going to entail different exercises and forms of power.

Many authors have approached conceptualizing power relations by classifying what different exercises of power look like and on what basis those exercises of power are successful. For example, French and Raven (1959) identify five ways power can be achieved: through referent power (power achieved or recognized through attraction or charisma), expert power (power achieved or recognized through relative expertise), reward power (the ability to provide incentives to achieve power), coercive power (the exercise of power through force), and legitimate power (power achieved through structural roles). Similarly, Nye (2008) states that power—as the ability to affect others to obtain outcomes—can be realized in three forms: coercion, inducements, or attraction (p. 94). While each of these taxonomies offers potentially fruitful inroads for understanding power as it exists in a relationship between a technology and a user, this dissertation turns to a different taxonomy of power built with an explicit consideration for how information technologies may impact power relations and built with Lukes and Foucault's conceptions of power in mind. Braman offers this taxonomy in her 2006 book *Change of State*.

Braman (2006) identifies power in four forms: instrumental power, structural power, symbolic power, and informational power. Instrumental power is power that "shapes human behaviors by manipulating the material world via physical force," structural power "shapes human behaviors by manipulating the world via rules and institutions," symbolic power "shapes human behaviors by manipulating the material, social, and symbolic world via ideas, words, and images," and informational power "shapes human behaviors by manipulating the informational bases of instrumental, structural, and symbolic power" (Braman, 2006, p. 25). While each of these forms of

power is distinct, Braman observes three important properties of these forms of power. First, they are often co-located. For example, in conflicts among state actors, a state may exercise both kinetic warfare (instrumental power) and propaganda campaigns (symbolic power) simultaneously towards achieving the same end. Second, the forms of power are often layered on each other and build on each other. For example, "smart weapons" layer informational power on top of conventional instrumental power, as these weapons can target specific individuals based on informational data (such as cellphone locations, GPS information) and can operate more independently of human intervention. Lastly, Braman observes that informational power can sometimes be a precondition for the exercise of other forms of power, going so far as to refer to informational power as "genetic" (2006, p. 26). For example, having a base of informational power can be a precondition for the exercise of instrumental power, such as when having the information on how to build a weapon (in the form of blueprints) is a precondition for building that weapon and then subsequently exercising the instrumental power inherent in the use of that weapon.[11]

Braman's taxonomy of the forms of power can inform how this dissertation conceptualizes the relationship between a technology and a user in three important ways. First and perhaps most obviously, Braman's taxonomy inherently highlights the role that the material artifact itself can play in plays of power in the relationship. For example, weapons (as a material artifacts) may serve as a form of instrumental power. Structural power may rely on material artifacts as a base for its processes, such as the reliance on electronic voting machines as part of the political process. Material artifacts such as the Internet may take the form of a medium through which propaganda circulates in the

[11] Informational power in this exact form has been a driver in some recent international conflicts, particularly ones involving the knowledge and informational basis associated with the development of nuclear weaponry.

enactment of symbolic power. Finally, databases can function as a base of informational power, as they may contain information, such as individual dossiers, that help a company target specific individuals for the exercise of other forms of power.[12]

Second, if, building on Foucault, the relationship between a user and a technology is conceptualized as involving relational power and acknowledges that each actor is both continuously undergoing and exercising power, Braman's framework further suggests that both actors may be undergoing and exercising differing forms of power in the relationship. That is to say, both a technology (as a sociotechnical system) and a user each maintain their own respective bases of instrumental, structural, symbolic, and informational power that inform the power-relations between the two. For instance, specific to the example of Twitter, Twitter maintains an informational archive of log data generated by each user, a base of informational power that it can use to determine, based on patterns of use, which users should be targeting for specific advertising campaigns (symbolic power). Users also maintain multiple forms of power in relation to the broader sociotechnical system (though admittedly, this power may not be as robust as the power of the sociotechnical system). For example, in the case of Twitter, a user may be capable of expressing symbolic power in petitioning Twitter's business owners for particular changes to platform. This leads directly to the third and most important observation for the purposes of this dissertation.

Third and lastly, Braman's observations about how informational power acts as a pre-condition or "genetic" factor for the exercise of other forms of power suggests considering how knowledge about a technology can act as a form of informational power, as that knowledge enables certain forms of action. This can occur in two ways. First,

[12] Such as the symbolic power inherent in advertising.

knowledge of a technology can help individuals realize affordances with that material artifact. The realization of these affordances may occur in the exercise of instrumental, structural, or symbolic power. Second, an individual's knowledge of a material technology can also enable certain forms of action or exercises of power in the relationship with the broader sociotechnical system. For example, knowing how a material technology operates at a technical level would better enable an individual to petition the sociotechnical system's designers for specific changes to a material artifact, if desired (an expression of symbolic power).

It is here that this chapter can directly integrate the work of Rogers into Braman's formulation of informational power. Rogers argues individuals may maintain different types of knowledge in relation to technologies, notably *awareness-knowledge, how-to knowledge,* and *principles-knowledge.* This chapter argues individuals who have more highly developed bases of *principles-knowledge* may be capable of realizing a great range of affordances in using a technology than those with simply *awareness-knowledge* or *how-to knowledge*. As the realization of affordances can occur in an exercise of instrumental, structural, and/or symbolic power, it is possible to say that these three forms of knowledge impact an individual's relative informational power. Further, these three forms of knowledge can also influence the range of power the individual can exercise in relation to the sociotechnical system. In summary, although though they do not exclusively constitute informational power, *awareness-knowledge, how-to knowledge,* and *principles-knowledge* of a technology all can contribute to the individual's relative informational power in the user-technology relationship.

Summarizing power in the user-technology relationship.

In conceptualizing the user-technology relationship, both the sociotechnical system and the user maintain a base of relational power that may be exercised in instrumental, structural, symbolic, or informational forms. On both sides of the relationship, informational power plays a critical role as "genetic" to the exercise of other forms of power. This dissertation, however, chooses to focus more heavily on user informational power within this relationship. For a user, knowledge about a material artifact can be a critical piece of informational power in their relationship with the technology. However, as Rogers observes, there are different types of knowledge about a material artifact. An individual who has developed *principles-knowledge*—as opposed to just *awareness-knowledge* or *how-to-knowledge*—of the material artifact is going to have a more widely developed base of informational power and may subsequently be better capable of exercising a greater range of power in this relationship.

However, as noted previously, an individual's knowledge of technologies within the world is mediated by their perceptions and experiences of the artifact and by the messages and media, as technological discourse, that they consume. This means that *principles-knowledge* and the informational power it enables has its genesis both internal and external to the user. Agents within a sociotechnical system, such as a technology's business purveyors, who are in a privileged position of generating technological discourse (an expression of symbolic power) and having it deployed through mass media thus have the opportunity to shape and influence users' knowledge of a technology, and thus the user's informational power. As a result, a user's informational power can be a contested site in the user-technology relationship.

The Emergent Theoretical Lens

To summarize the theoretical lens of this project and apply it to understanding the Twitter-user relationship, first, the relationship between individuals and technology develops within the broad context of technological diffusion. During Twitter's diffusion, individuals may have been exposed to the technological discourse around Twitter, may have watched others use the material artifact, or may have directly encountered the artifact. During this process of exposure, these individuals have built up (although not deterministically) different kinds of knowledge about Twitter: *awareness-knowledge*, *how-to knowledge*, and *principles-knowledge*. Rogers observes that these different forms of knowledge can influence and shape attitudes towards the technology, the use or non-use of a technology, and can be used as part of ongoing re-evaluation of technology use, if it has been adopted. Further, this knowledge may impact the range of affordances a user may realize with the material artifact in moments of use.

However, to better understand the relationship between Twitter and users, it is necessary to both take a broader view of what Twitter is and simultaneously recognize that there is nuance to users as a category. When users enter into a relationship with Twitter, they are not just in a relationship with a material artifact, but are also connected to a whole network of related production processes, actors, and variables that help manufacture and create the material artifact (such as Twitter's owners, internal programmers, developers), sets of skills and craft that helped to give rise to the material artifact, and the network of related practices, processes, and artifacts that have helped situate the material artifact in society (such as sales persons, advertisers, third-party developers, environmental factors, data resellers, rules and regulations, sets of laws, etc.). Users also should not be theorized as a singular homogenous mass. Instead, there are

different conceptualizations of users that may be applicable to understanding the Twitter-user relationship, such as differentiating users along the lines of different time of adoption, different use groups (including different non-user groups), and different groups as identified by other relevant social groups, such as Twitter's business owners.

Finally, in applying the discussion of power, the different forms of knowledge that these multitudes of users maintain can impact more than just technology adoption decisions. For a user, knowledge about the material technology of Twitter can be a critical piece of informational power. This is because knowledge about a material technology impacts the range of affordances a user can realize in the exercise of power and can shape the forms of power a user is capable of exercising in relation to the larger sociotechnical system. Not all knowledge leads equally to informational power however. As the discussion of Rogers's three forms of knowledge about technology highlights, different types of knowledge can lead to different abilities to realize affordances. Thus, these three types of knowledge will facilitate differing levels of informational power. For example, an individual who has *principles-knowledge*—as opposed to just *awareness-knowledge* or *how-to-knowledge*—of the material technology of Twitter may be better able to realize certain affordances of the material artifact during the exercise of power and may be better able to exercise power in relation to the sociotechnical system.

Individuals, however, develop *awareness-knowledge*, *how-to knowledge*, and *principles-knowledge* about technology from both their direct perception of a technology and through their consumption of technological discourse during the diffusion process. The messages created by organizations that purvey the technology can play an influential role in influencing an individual's *awareness-knowledge*, *how-to knowledge*, and

principles-knowledge regarding a technology. As a result, individual knowledge of technology, and thereby, informational power, is potentially (but not deterministically) influenced by the content of the technological discourse that generated by facets of the sociotechnical system.

With the exception of a few hypothetical examples, this review of the theoretical lens for the dissertation has not yet highlighted how specific technological features or specific business practices are collocated in this process. Instead, this conversation has remained broad and frequently hypothetical in nature. The next chapter of this dissertation addresses this by picking up the discussion of user power in the user-technology relationship and situating it within a contemporary history of social media websites, highlighting the distinct technical properties and business practices of social media websites that may impact this relationship. In doing so, the application of the framework presented in this chapter explores what many contemporary user-SMS power relationships look like and where critical scholarship has identified potential concerns regarding user power.

Chapter 3: Social Media, Information Flows, and User Power

Introduction

With the discussion of this dissertation's theoretical lens complete, this chapter focuses more directly on the relationship between users and a specific kind of technology: social media sites (SMSs). In order to describe this relationship, this chapter first defines "social media sites," a term used to refer to a genre of web-based technologies that have appeared within what O'Reilly (2005a) has called the recent history of "Web 2.0." Many SMSs technologies have common characteristics, including shared technical characteristics, common economic/network properties, and similar monetization practices. As part of this definition and broad descriptive account of SMSs, this chapter lays out the argument that users play a unique role within these spaces. On SMSs, users are not just consumers of the technology, but are also frequently the producers of the informational content that populates the platforms. Further, most for-profit SMSs are reliant on users' work in order to generate revenue, and, as a result, users' information creation and consumption practices are often directly tied to the profitability of these platforms. This chapter argues that, as a result, users' information creation and consumption practices are a critical component of many of today's popular SMS technologies.

After this brief introduction to SMS technologies and the role users play within them, the chapter turns to the issue of power in the user-SMS relationship. In reviewing the ways certain authors identify and describe power in the user-SMS relationship, two divergent strains of thought emerge: one that describes SMS technologies as power enhancing for users and another that suggests that users are often disempowered in their

relationships with SMSs. Of the latter, authors critical of user-SMS relationship dynamics often argue that users are put at a disadvantage in this relationship as a result of the business practices of the platforms, by specific technical configurations of the material technologies, and by the technological discourse that frequently surround the platforms. Further, many of these critics suggest that these impediments to user power result from social media sites' reliance on users as a source of labor and ultimately, revenue. In looking across these critiques of the user-SMS relationship, this chapter argues that one of the problems critics consistently identify is that users are inhibited in their development of knowledge and control over the information flows that exist on these platforms. However, this chapter also argues that this body of literature often only tacitly recognizes this issue as a problem of informational power.

The final section of this chapter argues for further empirical study into users' informational power on SMSs and suggests that the social media site Twitter makes for a timely and relevant space in which to pursue more detailed study into this area. To make this argument, this section introduces the Twitter, highlighting its current significance in the Web 2.0 landscape, and provides some anecdotal evidence as to why users' informational power in this space may be an issue. It also provides a brief overview of the extant research on Twitter and users, making note of a number of gaps in the way the user-Twitter relationship has been studied. Through the identification of these gaps, this final section demonstrates how this dissertation will make a unique and needed contribution to the scholarly work on informational power and on Twitter.

Social Media Sites

Social media sites (SMSs) are a genre of web-based technologies that have grown prominent during what some have referred to as the boom of the Web 2.0 Internet. There are many competing definitions for SMSs (Fuchs, 2014), though for this dissertation's purposes they can be broadly defined as, "a group of Internet-Based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user generated content" (Kaplan & Haenlein, 2010, p. 61). This user-generated content will also be referenced throughout this dissertation more generally as "user-generated information" or more simply "information." Examples of social media sites include places like YouTube, Facebook, Twitter, Reddit, Tumblr, Pinterest, Digg, Google Plus, Blogger, and hundreds (if not thousands) of others. Each of these sites vary in terms of the exact configuration of the material technology itself (such as the code, algorithms, protocols, and data structures of the site), the types and numbers of users on the sites, the types of informational content that can be shared within a platform (such as text, images, videos, etc.), the ownership status of the platform (such as publicly owned companies, privately held corporations), the governance of the technology (such as different terms of service and privacy policies), and the business models of the platform (van Dijck, 2013).

While SMSs are each distinct entities, many SMSs share common attributes. These shared attributes often stem from the sites' common grounding in the "ideological and technological foundations of Web 2.0" (to use Kaplan and Haenlein words). Therefore, the ideological and technological foundations of Web 2.0 are of relevance to understanding the user-SMS relationship as these common foundations shape and influence the configurations of both the material artifact the user interacts with and the

properties of the sociotechnical system the user is in relation with. Comprehensively accounting for the ideological foundations of Web 2.0 is beyond the scope of this chapter.[13] Instead, this chapter will focus on identifying a number of the common technological and economic foundations of Web 2.0 sites, of which, SMSs are a subset.

Common technical and economic foundations of Web 2.0 sites.

In 2004, Tim O'Reilly, the publisher of the O'Reilly technology books series, famously[14] promoted the term "Web 2.0" to describe what he saw as a new generation of web-based technologies. He defined Web 2.0 as:

> …the network as platform, spanning all connected devices; Web 2.0 applications are those that make the most of the intrinsic advantages of that platform: delivering software as a continually-updated service that gets better the more people use it, consuming and remixing data from multiple sources, including individual users, while providing their own data and services in a form that allows remixing by others, creating network effects through an "architecture of participation," and going beyond the page metaphor of Web 1.0 to deliver rich user experiences. (O'Reilly, 2005a)

Generally speaking, Web 2.0 sites include technologies such as blogs, wikis, social bookmarking sites, video-sharing websites, and social media sites. Andersen (2007), building from O'Reilly's definition, suggests that there are six important common technical and economic properties underlying Web 2.0 technologies: they are platforms for individual information production, they are platforms that harness the power of crowds, they are platforms that often manage large volumes of data generated by users, they are spaces built to be nearly-ubiquitously accessible platforms for participation and user contribution, they are platforms that benefit from network effects, and finally, there

[13] For an exploration of the ideological foundations of Web 2.0, see: Fuchs, 2010a, 2014; van Dijck & Nieborg, 2009.
[14] And some would argue infamously, see: Morozov, 2013.

is an openness element to the platforms. The following paragraphs describe each of these aspects in more detail.

Web 2.0 platforms, generally, do not have a staff that populates the content found within these sites. Instead, site users produce much of the information that makes up the Web 2.0 world. Andersen (2007) argues that within Web 2.0 environments, the historic capabilities of the "user" undergo a dramatic shift: users move from being the "passive" consumers of content to having the capabilities for "participatory" or "active" engagement.[15] Unlike television or radio users, Andersen argues Web 2.0 users can participate as both consumers of information and the producers of it. This particular aspect of Web 2.0 technologies has been described and debated heavily by scholars such as Jenkins and Deuze (2008), Bruns (2008), Shirky (2011a), Fuchs (2010b, 2014), and van Dijck (2009) (to name just a few). Neologisms such as "produser" (Bruns, 2008) and "prosumer"[16] (Quan-Haase & Young, 2010; Ritzer & Jurgenson, 2010) have been used to signify this relative shift in user capabilities. This chapter returns to discussing how power is involved in these "expanded capabilities" later.

Building from O'Reilly's comments that position Web 2.0 as a way of harnessing "the wisdom of the crowd," Andersen observes that many Web 2.0 platforms allow for the aggregation of collective intelligence or group work/production. Wikipedia, the collectively edited web-based encyclopedia, is frequently touted as an example of this sort of "crowd-sourced" effort. The term crowdsourcing was coined by *Wired* journalist

---

[15] There is, of course, a large body of work that argues that audiences are never actually "passive." Instead, many media theorists have suggested that different media facilitate different levels of engagement, sense-making and meaning-making; that audiences' ability to engage varies by individuals; and that engagement occurs even when an audience member cannot directly "talk back" through the media. See, for example, the differentiation made between hot and cool media by Marshall McLuhan (1964) and the work on differentiating audiences by television scholar John Fiske (1987).

[16] The term "prosumer" is generally credited to Toffler (1984).

Jeff Howe "to conceptualize a process of Web-based out-sourcing for the procurement of media content, small tasks, even solutions to scientific problems from the crowd gathered on the Internet" (Andersen, 2007, p. 16). While there may be a command and control structure of moderators and administrators that coordinate such crowd-sourced efforts, participation in most Web 2.0 spaces and crowd-sourced efforts therein is theoretically open to anyone who can connect. This particular facet of Web 2.0 technologies is also reliant on the positioning of the user as the producer of the informational content that makes up these spaces.

As users connect, browse, interact, and communicate (or "produse," to use Bruns terminology) within Web 2.0 spaces, they generate vast amounts of information, which the Web 2.0 platforms then house. While Andersen focuses more on how services such as Google and Amazon take advantage of such data rather than focusing on the statistics, Web 2.0 purveyors such as Facebook, Twitter and YouTube have reported gargantuan volumes of user-generated content being submitted to their sites. By the end of 2011, Facebook was the largest photo-sharing website on the Internet with over 250 million photos being uploaded by users *per day* (Horaczek, 2012). In 2013, Twitter received over 400 million messages (Tweets) from users *per day* (Tsukayama, 2013). In 2013, YouTube was receiving over 144000 hours of video from users *per day* (YouTube.com, 2013). These volumes are worthy of mention not just because of their magnitude, but also because there are particular economic benefits for Web 2.0 companies that can be derived from hosting large amounts of data. O'Reilly (2005b) suggests this when he states the value of Web 2.0 spaces is explicitly tied to the "scale and dynamism of the data [the technology] helps manage" (para. 15).

Andersen argues that one of the key characteristics of Web 2.0 sites is that they are vehicles for individual participation (as previously mentioned) *and* at the same time are also not dependent on users having specific technologies other than a device capable of accessing the Internet. Many Web 2.0 technologies offer ubiquitous access either directly through the user's web-browser, or more commonly today, through web-enabled mobile devices. Unlike traditional software that depends on the client's operating system and often, specific hardware, Web 2.0 technologies have been far more device independent. This has meant users frequently just need a device capable of connecting to the Internet in order to gain access to the platforms. According to Andersen, this allows a wider variety of individuals to participate in Web 2.0, as it reduces barriers traditionally associated with space, time, and place.

From an economic perspective, Web 2.0 platforms benefit from two types of network effects according to Andersen. The first is the traditional "network effect" in which the social and economic value of a communication network grows as new users are added. The greater the number of nodes in the network, the greater the overall value of the network, as the possibilities for connections between nodes grows in factorial size with every new node addition[17] (also called Metcalfe's Law). The second effect comes from "The Long Tail" phenomenon. Long-tail effects, as Andersen describes them, are essentially the ability to produce value from materials that are niche or of interest to very small populations. To illustrate this concept, consider a physical record store versus a digital record store like iTunes. A physical record store's shelf-space limits the catalog it can offer. It can only profit from what they can manage to fit on their shelves and are likely to therefore prioritize the most popularly selling materials in order to generate the

[17] This, of course, assumes that all nodes are valued as equal, and in different contexts it may not be.

most revenue. As a result, they will profit, but mostly likely only from what is popular with the majority of store-goers. iTunes, on the other hand, does not have these same kinds of space limitations and can essentially offer an unlimited catalog. Therefore, iTunes can benefit economically from selling both the most popular *and* the not as popular albums. They are able to extract profit from the "long-tail" of the niche, unpopular, or obscure. Web 2.0 platforms benefit from similar "long-tail" characteristics as the huge volumes of information that are created by users do not necessarily have to be curated, which more easily allows for niche information resources to thrive in these spaces and for value to be extracted from them. As a result, a platform like YouTube can benefit economically from hosting both popular videos such as Psy's 2012 "Gangnam Style" which has had billions of views since its uploading, as well as the 30 percent of all YouTube videos which have had less than 100 views (Frommer & Angelova, 2009).

Finally, Andersen argues that Web 2.0 platforms are "open" in a number of different senses of the term. He writes:

> The development of the Web has seen a wide range of legal, regulatory, political and cultural developments surrounding the control, access and rights of digital content. However, the Web has also always had a strong tradition of working in an open fashion and this is also a powerful force in Web 2.0: working with open standards, using open source software, making use of free data, re-using data and working in a spirit of open innovation. (p. 25).

It is worth noting however, that the openness that Andersen claims is inherent to Web 2.0 technologies is a highly contested notion and is a notion that this dissertation will challenge. This is discussed later as part of a review of critiques that have made of the user-SMS relationship. With some of the common technical and economic foundations of Web 2.0 having now been introduced, the next section of this chapter discusses another

common characteristic of Web 2.0 technologies which Andersen does not address: monetization practices.

Monetization of user-generated information.

While manufacturing is still an important aspect of the global economy, some have argued that countries such as the United States are now operating in a "knowledge economy" (Mokyr, 2004), an economic state where information, knowledge, and intellectual capital play a more predominant role in what makes up the overall economy. The shift towards a knowledge economy includes a change in the predominant commodities being produced. The critical political economists Hardt and Negri (2005) have argued that "immaterial goods" such as "ideas, knowledge, forms of communication and relationships" (p. 94) have become the dominant goods of production. Ideas, knowledge, forms of communication and relationships are also predominant outputs of users on most SMS sites and the businesses that run these spaces are frequently key players in extracting economic value from this information.

"Web 2.0 transforms the economics of knowledge-based businesses everywhere," writes Shuen (2008) in her book *Web 2.0: A Strategy Guide: Business Thinking and Strategies Behind Successful Web 2.0 Implementations* (p. 107). As much as O'Reilly (among others) argues that Web 2.0 is about specific features of material technologies, Web 2.0 is also heavily associated with certain ways that businesses generate value through the monetization of users' use of the technologies, and in particular, their information creation and consumption (Fuchs, 2014). For example, Scholz (2008) describes Web 2.0 as a phenomenon involving businesses profiting from "networked social production, amateur participation online, fan cultures, social

networking, podcasting, and collective intelligence" (p. 2). The commodification of user-generated information and information consumption has become strongly associated with the major players in the Web 2.0 world, most notably SMSs like Facebook and Twitter. Of the most popular Web 2.0 sites as measured by site-traffic (Alexa.com, 2015), almost all of these are run as for-profit businesses.[18] Facebook, YouTube, Twitter, Tumblr and Pinterest—all SMSs ranked in the top 50 of overall global web traffic—are all technologies put forward and run by companies that generate revenue from user generated information as well as users' consumption of information.

Many of the activities that users participate in within a SMS environment provide opportunities for the SMS business purveyor to generate revenue. There are two ways this generally occurs: through the sale of advertising displayed through the SMS to users and through the sale of access to information generated by users to third parties. Users' content creation and consumption practices are relevant to advertising as the information users generate serves as the "draw" for other users that will, in turn, consume both the informational content and the site's advertising content. Advertising is often displayed in close proximity to the user-generated content, sometimes blurring the line between the two. On Facebook, for example, advertisements appear on a scroll on the right hand side of the screen and within the users' timeline. N. Cohen (2008) describes this process generally, stating: "[b]y uploading photos, posting links, and inputting detailed information about social and cultural tastes, producer-consumers provide content that is used to generate traffic, which is then leveraged into advertising sales" (p. 7). The second way that some SMSs generate revenue is by taking the informational content that users create and then selling, sharing, or renting access to this data in its raw form or in

[18] The notable exception being Wikipedia.

aggregated user profiles to third parties. Twitter, for example, sells access to their "firehose," a pipeline of real-time user updates, and also through the company's Certified Products Program.

Based on these models of revenue generation, it is relatively straightforward to see how users' information creation and consumption practices can influence a platform's profitability. The more content users consume, the more they spend time also looking at advertising. The more content users create, the more that can be sold to third parties, or that can serve as content for other users to consume (and thus, those consumers spend more time looking at advertising). Users, however, generally do not share in any of the profit produced by these activities. Nick Bilton, a writer for the *New York Times*, glibly remarks on this reality in an article on Facebook's initial public offering of stock by stating:

> By my calculation, Mark Zuckerberg, Facebook's founder and chief executive, owes me about $50 . . . Facebook laid the foundation of the house and put in the plumbing, but we put up the walls, picked out the furniture, painted and hung photos, and invited everyone over for dinner parties. (2012, para. 1)

As N. Cohen (2008) puts it, "[b]usiness models based on a notion of the consumer as producer have allowed Web 2.0 applications to capitalize on time spent participating in communicative activity and information sharing" (p. 7). Users therefore play an incredibly important role for the overall profitability of these businesses, as they function not just the consumers of the technology in a more traditional sense, but also as an audience that consumes advertising, and a free labor source that generates content which is sold to third parties or has advertising sold against it (Scholz, 2008; Fuchs, 2014).

Summarizing SMSs.

SMSs can be considered as a "a group of Internet-Based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user generated content" (Kaplan & Haenlein, 2010, p. 61). These sites often vary in terms of the exact configuration of the material technology itself, the types and numbers of users on the sites, the informational content that can be shared within a platform, the ownership status and structure of the platform, the governance of the technology, and the business models of the platform (van Dijck, 2013). However, SMSs also frequently share common features or properties. For example, SMSs can be broadly described as being: platforms for individual information production, platforms that harness the power of crowds, platforms that often manage large volumes of data generated by users, spaces built to be nearly-ubiquitously accessible platforms for participation and user contribution, and spaces that benefit from network effects. Additionally, some (like Andersen) would argue that they have elements of "openness." Further, many of the businesses that run SMSs generate profits through the commodification of user information creation and consumption habits on these sites. As a result, individuals' use practices of the technologies can play a significant role in the commercial success and profitability of an SMS.

These technical, economic, and profit generating characteristics of SMS technologies inherently shape the relationship between users and the platforms. While this introduction to SMSs has tacitly touched on the user-SMS relationship, the user-SMS relationship is the explicit focus of the next section of this chapter. This discussion delves deeper into the relationship between users and SMSs and how power has been understood operating through it.

SMSs, Users, and Power

This section turns more specifically to the user-SMS relationship and the operation of power inside of it. While a number of scholars are optimistic about the kinds of individual capabilities engendered by these technologies, there have been a number of concerns raised regarding the disempowerment and exploitation that users face in their relationship with SMSs. Many of the concerns over diminished user power appear as a result of the monetization practices the rely on the commodification and alienation of user-generated information, as a result of certain technical configurations of the material technologies that make information flows on SMSs less visible to users, and as a result of technological discourses that surround the platforms that are often vague or unspecific about how information users create flows through the spaces. Fundamentally, many concerns over diminished user power appear to be connected to users' knowledge of, and control over, information flows on SMSs.

Empowerment/Exploitation.

A number of SMS purveyors, academics, and industry commentators have hailed SMSs as empowering for users. Many of these arguments follow the logic that individual power is increased through the use of these technologies as they allow users to engage in communication and join in "participation in media production and cultural expression" (Jarrett, 2008, para. 3). Through this individual expression, communication, and participation, users are better positioned to gain or maintain cultural and social capital (Ellison, Steinfield, & Lampe, 2007); undermine the authority of traditional media hierarchies and engage in participatory culture (Jenkins & Deuze, 2008); have the power to construct new identities, meet friends and colleagues and engage with strangers

(Albrechtslund, 2008); and even to become more active participants in governance (Shirky, 2011b). Further, many have argued that Web 2.0 technologies and SMSs in particular can engender group coordination and action (Shirky, 2011a; Tapscott & Williams, 2008), which, in turn, can facilitate users becoming part of what Castells (2009) refers to as networked counter-power. Even the language shift from "consumer" to "prosumer" seems to imply a relative shift in the capabilities and power of the user. As previously discussed, these sites are often described as "architectures of participation," rhetorically furthering the view that these technologies increase the power of users. After all how could "participation" be anything but empowering?

Jarrett (2008) argues that the celebration of the "newly empowered" user may in fact be premature. She writes: "Techniques of power which construct and promote this subject position serve to negate the hierarchy of traditional producer/consumer relations. Yet, this strategy can only function in relation to a producer/consumer power relation which remains … ultimately, unchanged" (para. 28). In essence, while users may gain access to expanded communication capabilities or the possibility of engaging in "participatory culture" through SMS platforms, this power is only gained through the imposition of a laborer/owner power dynamic in the user-SMS relationship. As van Dijck and Nieborg (2009) put it, while peer-production models appear to be replacing older top-down approaches and appear to yield more democratic informational structures, these exist "entirely inside commodity culture" (p. 855). Further, despite the arguments made by scholars such as Jenkins and Deuze (2008) that user empowerment in these spaces undermines the authority of traditional media hierarchies, users rarely have any measure of control over the information flows within these spaces (J. Cohen, 2012) and generally

do not share in any of the profit extracted from their labor (Terranova, 2000, 2004). Van Dijck and Nieborg (2009) provide an example of this logic when they state:

> Every user who contributes content – and for that matter, every passive spectator who clicks on user-generated content sites (such as YouTube) or social networking sites (such as Facebook) – provides valuable information about themselves and their preferred interests, yet they have no control whatsoever over what information is extracted from their clicking behavior and how this information is processed and disseminated. (p. 865)

Ritzer and Jurgenson (2010) argue that this type of relationship between user and SMSs like Facebook or Twitter represents a kind of "prosumer capitalism" in which "control and exploitation take on a different character than in the other forms of capitalism, [in which] there is a trend toward unpaid rather than paid labor and toward offering products at no cost" (p. 13).[19]

Scholars hailing from critical Marxist traditions argue that this is not inherently a new phenomenon, but is essentially old wine in a 2.0 bottle. They argue that the problem stems from age-old alienation and the exploitation of the laborer: the laboring class does not have power or control over the means of production (Petersen, 2008) and is exploited so the capitalist can gain surplus value (Fuchs, 2010b, 2014). The fact that the users do not own or control the means of production (essentially: the structure and code) of most SMS platforms means that they have little influence over what happens to the information they produce. Further, users are alienated from the informational product they created when this information is sold to third parties or is leveraged into advertising sales. As Fuchs (2014) puts it: "Corporate social media use capital accumulation models that are based on the exploitation of the unpaid labor of Internet users and on the commodification of user-generated data and data about user behavior that is sold as

[19] There is, of course, significant critique that this particular notion is somehow "new," see: Comor, 2011 and Fuchs, 2014.

commodity to advertisers" (p. 122). This alienation appears to be further reinforced by legal regimes that benefit businesses and leave little recourse for individuals to control flows of information to, from, and about themselves (J. Cohen, 2012). In this view, power relations between users and SMSs appear to be built on a foundation of labor alienation, exploitation, and ultimately domination (N. Cohen, 2008; Coté & Pybus, 2007; Fuchs, 2008, 2010b; Petersen, 2008). Clearly, this viewpoint does not match with a narrative of unbridled user empowerment. However, the outcome of user disempowerment may be achieved not just through alienation of users from the information they create or through legal and policy regimes that disenfranchise users, but also through the design of the material artifacts and the technological discourse that surrounds the sites.

Gaps in user knowledge of information flows as outcomes of design and discourse.

Despite Andersen's (2007) claim that one of the core characteristics of Web 2.0 sites is that they are open, this "openness" does not necessarily mean that the sites are transparent to users. Multiple studies have found that SMS users often maintain inaccurate, incomplete, or incorrect understandings of how the information they create moves throughout these platforms. For example, Acquisti and Gross (2006) found that almost a quarter of Facebook users misunderstood who could access their Facebook pages. Fuchs (2009) found that only 34 percent of Facebook users sampled in his study had "good or [a] high degree of knowledge on what Facebook is allowed to do with their data" (p. 113). In a test of user knowledge, Park (2013) found "more than 40% of the respondents misunderstood the most basic aspects of institutional data practices" (p. 224). Given their relationship to a technology's principles of operation, these gaps in

knowledge of information flows can be conceptualized as gaps in *principles-knowledge* of an SMS platform. As discussed in the previous chapter, gaps in *principles-knowledge* of a technology can manifest as diminished informational power for a user, resulting in a user being less able to realize certain technological affordances and less able to exercise instrumental, symbolic, structural, or informational power in relation to the material technology's sociotechnical system. When applied, gaps in *principles-knowledge* can contribute to negative consequences and outcomes for users. For example, in a qualitative study of user regret on Facebook, Wang et al. (2011) found that a lack of understanding of the ways Facebook makes information available to other users was often a contributing factor when users indicated that they had posted something on Facebook and later regretted it. As the previous chapter of this dissertation suggests, both a site's design and the discourse surrounding an SMS can perpetuate users' lack of *principles-knowledge* regarding a SMSs' information flows.

Design.

According to Langlois, McKelvey, Elmer and Werbin (2009), a common characteristic of commercial Web 2.0 platforms is that the ways that user-generated content are commodified are often kept invisible to users, specifically through the technical structures of these sites. Referring to this as "code politics," the authors state that many Web 2.0 purveyors make strategic design decisions to reduce user resistance, purposefully hiding how information flows through these systems and subsequently becomes commodified. Resistance, here, may mean different things. Principally, the term refers to a user potentially re-evaluating their use of a technology, choosing to use it differently, less, or to abandon it altogether. For the SMS purveyors, this may mean less

information creation and consumption, and therefore, less profit. Resistance can also mean other user expressions of power, such as instrumental, structural, symbolic, or informational power. For example, users creating a public awareness campaign to educate other users may be a form of "resistance."

To illustrate their point, Langlois et al. discuss Facebook's Beacon program, a controversial system that Facebook implemented that facilitated greater commodification of user-generated information. The authors argue that Beacon became controversial not simply because it involved the commodification of user-generated information, but because it became visible and known to users. Suddenly, through simply using the site, users were confronted with how the information they generated flowed through Facebook's platform to third parties, and how this flow was directly tied to the advertising the users experienced. Users petitioned Facebook for an immediate halt to the program and the program eventually became the basis of a lawsuit brought by a small number of users. Despite Facebook eventually putting a stop to that particular program, the "processes of commercialization… are still taking place on the Facebook platform" and, "these processes, however, increasingly take place at the back-end level and because they are invisible to users, they meet with less resistance" (2009, para. 17).

In engaging in "code-politics," in effect, platform purveyors are attempting to limit the development of users' *principles-knowledge* of particular information flows in order to culture particular information creation and consumption practices among users. The last chapter discussed how impeding development of a users' *principles-knowledge* can also impede that individual's informational power. As a result, the "code-politics" of a platform can impact and influence a users' informational power in the user-SMS

relationship. By reducing this type of user power, SMSs are essentially attempting to reduce "resistance" from users. In the case of Facebook's Beacon program, for users, the knowledge of how information flowed through Facebook to third parties in the process of commodification represented new *principles-knowledge* of the material technology, and hence, increased informational power. In gaining this informational power, some users began engaging new expressions of power in their relationship with Facebook, in the form of petitioning Facebook's business owners or in becoming part of the lawsuit against Facebook. Some users, unhappy with this newly revealed information flow, may have been less likely to generate informational content for the site, thereby threatening Facebook's ability to extract economic value from these users' labor. In order to rectify the situation, Facebook eventually removed the "Beacon" program. However, according to Langlois et al., Facebook also changed the visibility of other information flows in order to avoid stoking this same "resistance".

Stemming from this example, it appears that user perceptions and states of knowledge regarding SMS technologies are significant to its commercial success, as individuals make decisions about their use of SMS technologies based on these pieces of knowledge. It is in the economic interest of these companies to promote states of user knowledge that are conducive to the users' participation and labor, directing users away from states of knowledge that might give them reason to behave differently (such as exposing them to information flows that might give them qualms). As a result, SMSs may have strong motivations for structuring their material technologies in ways that promote *awareness-knowledge* or *how-to knowledge* of a SMS technology, but may have markedly less motivation for promoting *principles-knowledge* of a technology where that

principles-knowledge deals with information flows that could provoke or invite resistance. As a result, the purveyors of these technologies may—intentionally or unintentionally—inhibit users' informational power through the material configurations of the technologies.

Discourse.

The work of van Dijck and Nieborg (2009) suggests that users' difficulty in developing knowledge about the information flows within SMSs may be perpetuated not only by the structuring/code politics of a site, but also by the technological discourse that surrounds these spaces. In analyzing a number of Web 2.0 business manifestos, the two observe:

> Web 2.0 manifestos … typically *do not provide any technological details about how various sites render profitable business models* … they focus on the emancipation of consumers into users and co-creators, *rather than on the technical details concerning how these sites turn a profit* [emphasis added]. (p. 866)

This is to say: the language that Web 2.0 purveyors use to describe their technologies may not contribute much to the development of *principles-knowledge* of information flows within these platforms. Gillespie (2010) similarly argues, "Online content providers such as YouTube are carefully positioning themselves to users, clients, advertisers, and policymakers, making strategic claims as to what they do and do not do, and how their place in the information landscape should be understood" (para. 1). The language that a Web 2.0 purveyor chooses to position itself is critically important when seen as factor facilitating a user's adoption and use of an SMS technology.

Besides mission statements and business manifestos, privacy policies and governing documents are another way that websites communicate what they do with the

information they collect from users. According to Jensen and Potts (2004), privacy policies "are meant to inform consumers about business and privacy practices and serve as a basis for decision making for consumers" (p. 471). However, much like the material technologies themselves, these statements are often less than transparent. Cranor (2003) argues: "read-ability experts have found that comprehending privacy policies typically requires college-level reading skills" (p. 50). Further, Fernback and Papacharissi (2007) argue that, while privacy policies may describe the general kinds of information a company collects, the language used to describe information collection, use, commodification, and sharing is often quite broad. According to Fernback and Papacharissi, many privacy policies are also vague or unspecific about the particular third parties that user-generated information is shared with or sold to.

While Jensen and Potts (2004) argue that privacy policies are meant to inform consumers about what businesses collect and do with individuals' information, Fernback and Papacharissi (2007) instead suggest that "Privacy statements, crafted by staff attorneys, are written to coincide with business models so that firms may maximize the ability to profit from information that they capture" (p. 719). Rather than being technological discourse designed for benefit of users, Fernback and Papacharissi argue that these documents are constructed so that, in the event of a complaint, the company can be absolved of legal responsibility. As a result, the documents are more likely to be written in legal language that maximizes the flexibility afforded to the businesses that run these sites than in language that clearly communicates to users.

Taken together with the work of Langlois, McKelvey, Elmer and Werbin (2009), a rather bleak picture of the transparency of information flows on SMS platforms

emerges. When the flows of information through these spaces are not perceivable through the structure of the sites and are not easily discoverable through the technological discourse around these sites, users are put at a significant disadvantage in their abilities to build *principles-knowledge* of information flows. As a result, users are put at a de facto disadvantage in their ability to develop informational power in relation to the SMS. This impediment to the development of user power can have significant impacts for the individual.

Why knowledge of information flows is important for users.

So far, this chapter has highlighted ways that SMSs both engender positive outcomes for user capabilities and power, as well as the ways SMS users are disempowered. Taking a more critical view, this review has highlighted ways users' informational power can be negatively impacted by "code politics" that make information flows on SMSs less visible and by technological discourses in business manifestos and privacy policies that frequently fail to depict the information flows present in these spaces in detail. These observations demonstrate how a user's *principles-knowledge* of the information flows of an SMS, as a form of individual informational power, can be a site of contention within the user-SMS relationship. This occurs particularly as the businesses that operate many of these platforms look to mitigate "resistance" from users so as to maximize profit generation. However, so far, this chapter has only tacitly addressed why a lack of informational power is problematic for users. Next, the chapter addresses reasons why a user's knowledge of information flows (and hence, informational power) is important.

There are two general, but related, areas of concern regarding users' knowledge and power in relation to information flows. The first is a set of "structural" concerns about diminishing user power. For example, van Dijck and Neiborg (2009), argue that user disempowerment is problematic as it results in alienation, where users have little to no control over the information they themselves created. Terranova (2004) argues that user alienation and disempowerment is problematic as it shuts-out users from controlling the means of production and from access to the profits generated through their labor. Puschmann and Burgess (2013) argue that "code-politics" is problematic for users because it shuts them out of shaping the "emerging data market" according to their interests. More broadly, J. Cohen (2012) argues that non-transparency of information networks can be a detriment to human flourishing. These examples reveal how users are put at a structural disadvantage in their relationship with the sociotechnical system the SMS is part of.

The second reason why barriers to user power can be problematic can be thought of as "application/decision making" problems. This line of thought argues that barriers that prevent users from developing knowledge of how information flows through SMSs like Facebook and Twitter can be problematic for users when users apply their incomplete knowledge and make decisions about a technology's use. For example, as previously mentioned, Wang (2011) found that confusion regarding information flows on Facebook was frequently a contributing factor when users regretted posting content. Similarly, Baker (2011) expresses a concern that without knowledge of how information flows through Twitter, users are less able to make informed decisions about what they post. Other scholars, such as Mahmood and Desmedt (2012) argue that in order to

enhance users' abilities to manage their privacy, information flows in SMSs need to be made clearly visible to users. By making information flows clearly visible, users can have more access to knowledge which allows them to better assess the risks associated with information disclosure and participation in SMS spaces. Hull, Lipford, and Latulipe (2011) make a similar observation about Facebook, arguing from an ethics perspective that users' privacy is put at risk by a lack of transparency regarding information flows within the site. The authors conclude, "Facebook needs to do a better job of making the flows of information on the site transparent to users" (p. 300). Although Hull, Lipford, and Latulipe do not empirically study users' understandings of information flows, the authors go so far as to say that Facebook has a tendency to encourage mistakes in users' risk perception by not clearly stating how information is made available to other users and to third party companies. These examples reveal how users are put at a disadvantage as they apply their knowledge of information flows to decision making.

Regardless of which area of concern is more compelling, authors working in both streams acknowledge that a lack of knowledge regarding SMS information flows is problematic for users. Not only do users often have incomplete, inaccurate, or otherwise incorrect understandings of information flows in Web 2.0 and SMS spaces, but their development of more robust understanding can be hindered by the sites' structures and by the discourse surrounding them. These become barriers inhibiting the development of individual users' informational power, leading to a number negative outcomes for users, both structural and in application. However, there are a number of outstanding questions highlighted by this body of work that have yet to be addressed. First, while many scholars provide excellent analysis of code politics in SMSs, these investigations are

rarely paired with empirical user studies that trace individual users' knowledge of the platforms. Second, while there has been some analysis of the technological discourse around SMSs, there has been little systematic evaluation of the ways that the discourse generated by SMSs depicts informational flows or how this might relate to users' *principles-knowledge* of informational flows on the platforms. Finally, and perhaps most importantly, there are few studies systematically approaching SMS users' knowledge of information flows on these sites in relation to how this can impact their relative power. As a result, the interconnections and tripartite operation of SMS users' knowledge of information flows, technological discourse, and informational power in the user-SMS relationship have not been thoroughly studied. Next, this chapter will argue that a coherent study that explores these outstanding issues in tandem is needed and will propose Twitter as a timely SMS in which to pursue further study.

Twitter as a Case

Proceeding from this gap in the scholarly literature, this section argues that the social media platform Twitter is an apt space to investigate the state of users' informational power vis-à-vis their *principles-knowledge* of information flows, and how the technological discourse around this platform might influence this informational power. First, this section briefly introduces Twitter, explaining what it is, the site's history, and its significance in the contemporary SMS landscape socially, culturally, politically, and economically. Next, this section surveys previous research on the user-Twitter relationship and identifies some of the remaining gaps in this literature; particularly gaps related to users' knowledge of the information flows on Twitter and of the technological discourse surrounding the platform. The chapter concludes by staking

out a set of research questions regarding users' knowledge of informational flows and technological discourse, questions that focus on more clearly articulating the state of users' informational power in the user-Twitter relationship.

What is Twitter?

Founded in 2006, Twitter is a social media site that primarily facilitates the exchange of short-form (140 character) messages; a type of interactive format sometimes referred to as microblogging (Java, Song, Finin, & Tseng, 2007). For years on Twitter.com, "What are you doing right now?" was the compelling question greeting site visitors. Historically, Twitter's prompt has served as one of the most prominently displayed messages on the Twitter homepage, an instruction suggesting how individuals should use the service and stressing the importance of the present and current on the site. Of course, as users of Twitter are likely aware, not everyone chose to answer this question. Barb Dybwad (2009), a writer for online tech news site *Mashable* notes, "the official question is largely ignored by those who have found myriad ways to share pretty much anything they wanted, be it information, relationships, entertainment, citizen journalism, and beyond" (para. 2).

This sharing of information, relationships, entertainment, journalism, and beyond has made Twitter a phenomenon in the world of Web 2.0 technologies. In late 2013, traffic on Twitter was clocked at roughly 500 million Tweets per day and the service had an estimated 200 million monthly users (Kim, 2013). In an interview with Liane Hansen of National Public Radio, Andy Carvin, a strategist for National Public Radio's Social Media Desk, succinctly summarized Twitter this way:

> Twitter in many ways has become the pulse of what's going on online right now. Because it's a real-time conversation that anyone can chime into at any given

> point, it's 24-7. And so when something happens somewhere in the world you're almost guaranteed that people will be talking about it or even witnessing it as it happens, whether it's protests and revolution in Kyrgyzstan to people talking about the ham sandwich they just ate and everything in between. (As quoted in Hansen, 2010)

Twitter's large user base, its hyper focus on real-time communication, and the site's massive amount of informational content has made it an attractive site to a number of different users for a multitude of purposes. While Twitter has become a vehicle for communicating the mundane activities of a user's daily social life (boyd, 2009; Miller, 2008), the platform and the information made available through it have also become significant for cultural, political, and economic ends. For example, Twitter has been used by political activists to get their messages out to the public and for coordinating protest activities during events such as the 2007 Nigerian Election protests (Ifukor, 2010) the 2008-2009 Iranian protests (Burns & Eltham, 2009; Grossman, 2009), and the Occupy protest movement (Juris, 2012; Thorson et al., 2013). Politicians have used Twitter to engage constituents and to rally support around specific political issues and during elections (Golbeck, Grimes, & Rogers, 2010; Grant, Moon, & Busby Grant, 2010; Hong & Nadler, 2011; Jackson & Lilleker, 2011). Information on Twitter has been used by brands wishing to understand up to the moment sentiment about their products (Jansen et al., 2009). Tweets have been used by academics and by governments for detecting real-time events such as earthquakes (Earle et al., 2011; Sakaki et al., 2010), natural disasters (Bakshi, 2011) and to monitor disease propagation (Achrekar et al., 2011; Lampos et al., 2010; Signorini et al., 2011). Sentiment in Tweets has even been used to try to predict short-term performance of the stock market (Bollen et al., 2011; Sprenger & Welpe, 2011; X. Zhang et al., 2011). As Twitter's user-base has created a deluge of up-to-the

second information and activity, Twitter has become a prominent means for individuals, academics, businesses, governments, news organizations, and others to access and interact with the *zeitgeist* of the Internet.

The Library of Congress confirmed the historical value of this body of messages when it announced in 2010 that it was partnering with Twitter to store all public Tweets, ever. Twitter had agreed to donate a gift copy of its entire archive to the Library of Congress with the additional promise of all public future Tweets. The Library of Congress has since justified accepting this acquisition by stating:

> Twitter is a new kind of collection for the Library of Congress, but an important one to its mission of serving both Congress and the public. As society turns to social media as a primary method of communication and creative expression, social media is supplementing and in some cases supplanting letters, journals, serial publications and other sources routinely collected by research libraries. Archiving and preserving outlets such as Twitter will enable future researchers access to a fuller picture of today's cultural norms, dialogue, trends and events to inform scholarship, the legislative process, new works of authorship, education and other purposes. (Allen, 2013, para. 6)

The Library of Congress was not alone in announcing its partnership with Twitter. Twitter, in its own press release, explained that Tweets have "become part of significant global events around the world" (Stone, 2010b, para. 2), and that, "[a] tiny percentage of accounts are protected, but most of these Tweets are created with the intent that they will be publicly available" (para. 2). However, this presumption of user intent may be somewhat problematic. As highlighted in the previous section, users often have incomplete understandings of how information flows through Web 2.0 platforms, as a sites' code-politics and surrounding technological discourse can sometimes impede the development of users' *principles-knowledge* of information flows. Impediments to the development of *principles-knowledge* also inhibit the development of users'

informational power. Furthermore, the companies that operate Web 2.0 sites may be motivated to structure their sites or construct their discourse in ways that do not fully reflect the information flows of the platform so as to not promote "resistance" from users. What is not yet apparent is whether or not these same tensions can be seen between Twitter and Twitter users.

As noted in Chapter 1, despite the potential research value of a Library of Congress archive, some users of Twitter were not pleased with the announcement. Comments from Twitter users on the Library of Congress' blog indicate surprise and frustration regarding the seemingly newfound permanence of Tweets. As was made obvious by the fact that Twitter gave the Library this archive, however, Tweets have never been fleeting. Twitter has always kept the Tweets sent through its system. However, some users may not have anticipated that information would ever flow from Twitter to a recipient like the Library of Congress. Manifested in the comments left on the Library of Congress announcement, there appears to be a gap between these users' *principles-knowledge* of information flows on Twitter and how Twitter actually managed the messages sent through the service.

From this anecdote about the Library of Congress Twitter archive, let us imagine a hypothetical user. Based on their incorrect perceptions about the way that information flows on Twitter, this user may have sent a message through the system that they did not expect to be archived, let alone archived in the Library of Congress. Perhaps this message contained embarrassing, personal, or otherwise sensitive information. Perhaps this message was not even particularly salient or embarrassing at the time, but in a context five years from now, becomes relevant at a job-interview. This imagined user who

operated under a false perception of Twitter's informational flows (and who therefore had a weakened base of informational power), might not have posted this message on Twitter had they understood these flows more accurately. Is this hypothetical example something that is actually occurring systematically on Twitter though? Without further study, it is difficult to say. What it is possible to say, however, is that Twitter is an excellent candidate for studying users' informational power vis-à-vis their knowledge of information flows, not just because of the site's current social, cultural, political, and economic uses and significance, but because—despite the misunderstandings highlighted within the Library of Congress Twitter archive comments—it is touted as a transparent and uncomplicated platform in comparison to other such Web 2.0 platforms (Bruns, Burgess, Crawford, & Shaw, 2012).

Relevant Research on Twitter.

While hundreds of academic studies have been undertaken with the use of data from Twitter (Zimmer & Proferes, 2014), there are markedly fewer studies that focus on users' understandings of the information flows, analyze Twitter's technological discourse, or address issues of user power on the site. What this chapter presents in the following review is not an exhaustive accounting of all the existing work on Twitter that touches these, but is instead a selected highlighting of important and salient work that points the way for further study.

User knowledge of the Twitter platform.

While there are no studies directly assessing user understandings of information flows on Twitter, there are a few studies that broadly explore users' beliefs about the site. In their 2011 qualitative study of Twitter users, Marwick and boyd explored how Twitter

users imagine the audiences for their messages (conceptualized here as "potential receivers" of information flows). They found that users engage in a number of different cognitive strategies to envision their audiences before constructing Tweets. For example, many users indicated that they conceptualize close personal friends, persons that resemble themselves, and communities of interest as the potential recipients of messages. The authors found that some users actively self-censored themselves based on the fear that their employers or community authority figures, such as parents, would find Tweets. However, Marwick and boyd do not discuss whether users conceptualize Twitter's commercial partners as receivers of these messages. In fact, the authors found "[p]eople with few followers, who use the site for reasons other than self-promotion, generally see Twitter as a personal space where spam, advertising, and marketing are unwelcome" (p. 11). This does not necessarily mean that individual users are unaware that Twitter sells access to the data that users generated, or that users are unaware of how Twitter structures its information flows, just that users do may not actively consider those recipients of information as part of their "audience".

In their article, "The Tweet smell of celebrity success: Explaining variation in Twitter adoption among a diverse group of young adults" Hargittai and Litt (2011) explore the attributes of Twitter users and non-users through a survey of over 500 undergraduate students at the University of Illinois, Chicago. The pair noted that, despite 62% of the students not being Twitter users, only 2% had never heard of Twitter (indicating a high level of *awareness-knowledge* among this population). The pair also found that adoption among the sampled population is not uniform and that there are several notable characteristics of those who adopt the technology. The authors observe

that technological skill levels are correlated with Twitter adoption, with "those with higher skills… [were] more likely to use the service" (p. 835). Unfortunately, other than the participant's general awareness of the technology's existence, their general technical skills, and their use/non-use of the platform, the authors do not go any deeper into users' knowledge of the Twitter platform.

González and Juárez (2013) have, perhaps, most directly tackled the topic of users' conceptualizations of the Twitter platform. The pair elicited the mental models that a group of 30 undergraduates used to conceptualize the Twitter platform and then correlated these models with users' success at completing basic tasks on Twitter, such as retweeting a Tweet and sending a direct message. The students' mental models fell into one of three categories: an "analog" mental model, a "technical" mental model, or a "conceptual" mental model. In the analog mental model, participants indicated they understood Twitter by analogy with other activities (for example, stating that tweeting is like talking or that the timeline is like a chalkboard). In the "technical" mental model, participants indicated that they understood Twitter more technically, for example, as a set of massive databases connected together with the devices that individual users tweeted from. Finally, in the conceptual mental model, participants "only wrote the concepts and their relations" or "used graphic elements of the system of screenshots to describe them" (p. 9). González and Juárez found that, regardless of mental model, users were able to complete basic tasks on the site, though newer users had a tendency to take a longer time in accomplishing the task. The work of González and Juárez highlights the fact that (as discussed in Chapter 2 of this dissertation) users need not have a complete or even accurate picture of how the technology works in order to be able to use it. Unfortunately,

their work also has a significant number of limitations. In addition to limitations based on sample size, the project focused more specifically on users' ability to complete tasks such as sending a tweet, rather than going into more detail regarding the accuracy of the participants' mental models. This analysis also lacks a broader connection to issues of user power.

Twitter and discourse.

While there are a number of studies that have explored discourse found within user-generated Tweets, far fewer have explored the technological discourse that surrounds Twitter. Arceneaux and Weiss (2010) trace the early discourse about Twitter as it played out in the popular media from 2006-2009. The authors found that a majority of the newspaper articles about Twitter during this period address what Twitter is, the novelty of the medium, and focus predominantly on the Twitter platform's brevity and speed (i.e., the fact that messages on Twitter are short and are disseminated instantly). While these are two important pieces of information that could spur the development of *principles-knowledge*, more robust descriptions of the information flows on Twitter were generally not present in this work. The authors found that there were a minority of articles that covered Twitter in a critical manner, however, rather than detailing how the technology functions most critical coverage of Twitter focused on the problem of "information overload" as a social phenomenon (e.g., it is hard to keep with all of the messages), "acceptable practices" (e.g., what informational content not to tweet) and the problems of impersonation that could occur on the site. Interestingly, at the time, Twitter did not monetize user-generated information. Arceneaux and Weiss note that many of the articles discussing Twitter repeatedly expressed concern over Twitter's apparent lack of a

business model. Unfortunately, this study does not explore Twitter's own organizational rhetoric, though it does provide a number of useful insights into how the technology was being described in the mass media during these early years.

While Arceneaux and Weiss focus exclusively on discourse in the public media, Baker (2011) takes up the question of Twitter's own messaging in regards to the Library of Congress archive. In arguing about the ethics of the Library of Congress archive, she describes Twitter's documentation of the archive's existence to users as "not reassuring" (p. 10). In her 2011 examination of Twitter's Terms of Service and Privacy Policies, Baker finds not a single mention of this archive, and highlights the fact that Twitter's messaging to users does not actively disclose that this information flow to the Library of Congress exists. She concludes:

> Although it is not providing users with incorrect or false information, Twitter is capable of disclosing the Library of Congress Twitter archive in a more straightforward way. Explicit references to the archive institution and the restrictions placed on the archive would educate users and enable them to make more informed decisions about what they post. (p. 11)

It should be noted that Twitter did begin mentioning the Library of Congress archive in a mid-2012 Privacy Policy update, but that Twitter essentially went two years before disclosing this information flow as part of their governance documents. While useful, missing from Baker's analysis is an empirical study of users' states of knowledge regarding this particular information flow.

User power.

Users' power in relation to Twitter has not been written about extensively, but it does appear either tacitly or explicitly as the subject of a few studies. For example, in tracing the technical evolution of Twitter during its first three years of existence, Siles

(2013) argues that the material configuration of Twitter was greatly influenced by "feedback loops" (a phrase borrowed from Hayles [2010]) that ran between Twitter's users and Twitter's developers. To illustrate his point, Siles describes how users first began using the "@" symbol to address each other in conversation and how this communication convention then became formalized in the system when Twitter engineers took notice and altered the protocols of the platform. Even though the final form of the reply and mention protocols were ultimately determined by Twitter's engineers, Siles argues that the shape of Twitter today has been greatly influenced by users. Although Siles does not provide a formal analysis of power in the user-Twitter relationship, he does ascribe a heightened level of influence in how this technology emerged and crystallized to particular users. Siles appears to suggest that not all users are totally cut off from influencing the means of production. Upon second glance, however, what makes Siles' work less optimistic is that many of the iterative feedback loops that Siles identifies as pertinent to the shaping of Twitter existed primarily between Twitter developers and either users who were their friends or users who were also engineers in Silicon Valley. Specific users seem to have this power to influence, and that influence appears dependent on social positioning and access. Further, the feedback loops that Siles identifies were primarily present during Twitter's early years when its user base was far smaller. Given the growth of Twitter's user-base since 2006, it is difficult to imagine that the same proportionality of user influence exists today.

Van Dijck (2011) provides a similar analysis of the early years of Twitter's infrastructure in her article, "Tracing Twitter: The rise of a microblogging platform." In this work, van Dijck addresses the question of how Twitter and the practice of

microblogging evolved during the platform's first five years of existence and how Twitter came to have a dominant meaning. Rather than addressing these questions through the lens of feedback loops, van Dijck relies on the social-constructivist concept of interpretive flexibility to understand the "mutual shaping of technology and users" (p. 2). Unlike Siles, van Dijck specifically attempts to expose the plays of power in the process of meaning stabilization. Instead of framing interpretive flexibility as co-creation or a collaborative process where everyone is working towards the same end, she instead observes it playing out on Twitter as a confrontational process where various actors are engaged in a struggle over power. She writes that ultimately, "Twitter's meaning as a tool and service will be as much the result of conscious steering by its owners as of accepting and/or resisting such steering by users…" (p. 20). Although van Dijck observes that users were highly influential in the process of meaning stabilization early on in Twitter's existence, she notes that by 2011, "the meaning of microblogging is still flexible but less so than five years ago" (p. 19). Van Dijck goes on to observe that since 2010, as Twitter's monetization practices have increased, Twitter's founders have more at stake in fixing the technology's meaning: they must now also do it in a way that does not alienate their users. For example, "if users resent promoted Tweets [among other advertising] in their personal content stream, they may instantly quit Twitter" (p. 18). This is similar to the observation made by Scholz (2008) and Langlois et al. (2009) that Web 2.0 and SMS purveyors engage in a careful act to extract profit from user labor without driving that labor away or provoking resistance. While van Dijck traces many of the prevalent interests that have shaped Twitter's meaning as a sociotechnical object and how power is

part of this process, this analysis provides more of a historical account and lacks a focus on of how discourse and user knowledge are collocated in this process.

In her 2012 dissertation, *Programmed sociality: A software studies perspective on social networking sites*, Bucher focuses a chapter on the development of the Twitter application programming interface (API) ecosystem. In this chapter, she analyzes the API interfaces, interviews API third party developers and users, and examines Twitter's organizational rhetoric around the APIs to explore the ways in which meaning and possibilities for action are constructed in relation to the APIs. She writes that the APIs, "shape, control, and enable practices of sharing, transmission, and innovation in multiple ways… APIs thereby have very real material effects on end users. On the other hand, APIs regulate and restrict the same flow of data and information that they enable" (2012, p. 191). Ultimately, Bucher argues that the APIs are a space where there is an intense struggle over structural power among third-party developers and Twitter itself. She notes that Twitter uses its governing documents; its ability to control the interface, algorithm, and protocols; and its own discourse to set essentially the rules for information flows that ultimately underpins freedom and control for developers. As a result, third-party developers have had progressively diminished power in their relationship with Twitter. Bucher establishes an excellent argument about how Twitter deploys discourse to shape power in the relationship with its third-party developers, but developers' knowledge and informational power are not specifically studied in depth as part of this analysis.

Puschmann and Burgess (2013), building on the work of Gillespie (2010), evaluate Twitter's technical infrastructure through the lens of a version of code-politics they call platform politics. In this critical evaluation, the pair argues that non-developer

and non-advertising users on Twitter lack true control over the information they create, and conclude:

> In the current state, the ability of individual users to effectively interact with "their" Twitter data hinges on their ability to use the API, and on their understanding of its technical constraints. Beyond the technical know-how that is required to interact with the API, issues of scale arise: the Streaming API's approach to broadcasting data as it is posted to Twitter requires a very robust infrastructure as an endpoint for capturing information ... It follows that only corporate actors and regulators— who possess both the intellectual and financial resources to succeed in this race—can afford to participate, and that the emerging data market will be shaped according to their interests. End-users (both private individuals and non-profit institutions) are without a place in it, except in the role of passive producers of data. (p. 11)

Andrejevic (2013) makes a similar observation regarding Twitter's infrastructure, noting "That anyone who wishes can use Twitter to express him- or herself is something very different from anyone being able to access and mine Twitter's 'firehose'" (p. 181). However, there appear to be two separate issues of power at play in the observations of these scholars. First, both sets of authors seem to be observing that access itself is an inherent issue of power in the user-Twitter relationship. The inability of some users to access certain interfaces, such as the Streaming APIs, disempowers those users. However, Puschmann and Burgess *also* inherently argue that, without *knowledge* of the APIs, users may additionally find themselves in the position of being passive producers. Unfortunately, neither Puschmann and Burgess, nor Andrejevic, empirically explore Twitter users' knowledge of the information flows relative to the platform. Also absent from these critical perspectives are an account of how the discourse around Twitter might play a role in shaping knowledge of the information flows.

The work that remains to be done.

It is possible to identify a number of gaps in the research on Twitter users' understandings of the platform, the discourse surrounding Twitter, and users' power. Previous research has explored Twitter use and users' general familiarity with the site, how users conceptualize audiences on Twitter, and users' mental models of the platform. However, this work does not include in-depth exploration of Twitter users' *principles-knowledge* of the information flows of the platform. Additionally, the more empirical investigations into users' understandings and knowledge of the platform that do exist are rarely accompanied by an analysis of power. Next, while there has been some analysis of the discourse around Twitter, a broad and systematic evaluation of Twitter's organizational rhetoric regarding the platform's informational flows is missing. This discourse is potentially important as it may impact users' knowledge regarding the information flows of the platform. Third, while there are some scholars taking a critical look at issues of user power in relation to Twitter, these studies frequently lack more empirical components relating to user-knowledge. Fourth, the studies that have explored issues of power have tended to focus on issues of structural power rather than informational power. Finally, while there has been some focus on the API interfaces as a particular site of struggle in the user-Twitter relationship (developer-Twitter more specifically), other components, such as the general web-interface, have not been given as much critical attention.

Filling these gaps is an important step in understanding users' informational power in relation to the Twitter SMS. However, equally important is understanding how these elements of user knowledge of information flows, discourse, and power work together in tandem. It is with this need in mind that this dissertation asks as its primary

research questions: In the user-Twitter relationship, what is the state of Twitter users' informational power in regards to the informational flows of the platform? What *principles-knowledge* of information flows do users have and how does the technological discourse surrounding the site created by Twitter's business purveyors describe information flows and potentially impact users' *principles-knowledge* (and hence informational power)?

In order to answer these questions—and in order to explore what users know about information flows on Twitter, and in order to assess how Twitter's organizational rhetoric describes the platform's informational flows—this dissertation must first outline and describe how user-generated information actually flows through Twitter. To map the information flows of Twitter (and in order to also explain, to a lesser degree, why it flows in the way it does), the next chapter of this dissertation deconstructs Twitter as a sociotechnical object. The chapter does this using an analytical framework provided by van Dijck (2013) in *The Culture of Connectivity,* her critical history of social media. Van Dijck's framework suggests that an SMS can be understood through its technical structure (its interfaces, algorithms, protocols, defaults, and data and metadata structuring), the kinds of users that exist on Twitter, the kinds of informational content found on the platform, the SMS's business models for the platform, the SMS's ownership status, and the SMS's governance. The next chapter's analysis serves as the basis for both a study exploring users' knowledge of information flows on Twitter and a discourse analysis that explores how the company describes its information flows. This analysis will enable this dissertation to better explore user knowledge of informational flows and

the technological discourse around Twitter, contributing towards a clearer articulation of informational power in the user-Twitter relationship.

Chapter 4: Deconstructing Information Flows on Twitter

Introduction

To recap the argument of this dissertation so far, Chapter 2 began by positing a theoretical lens for conceptualizing the relationship between a user and a given technology. Chapter 2 argues that: the relationship between a user and a technology develops as the technology diffuses throughout society, that technologies can be conceived of as complex sociotechnical systems, and that the term "users" can refer to a number of different types and stratifications of individuals. Chapter 2 continues, arguing that the relationship between a user and a technology (being conceptualized as a sociotechnical system) often involves the negotiation of power in multiple forms. A particularly important form of power in the user-technology relationship is what Braman (2006) identifies as informational power, a form of power that often serves as the base or precondition for instrumental, structural, or symbolic power. An individual's informational power can be impacted and influenced by the knowledge that individual maintains regarding how a given technology functions (referred to as *principles-knowledge*). The development of *principles-knowledge* of a technology can be influenced by factors external to the individual, such as the material design of the technology and the technological discourse that surrounds the artifact. Finally, Chapter 2 concludes by arguing that, despite the influence on individual's *principles-knowledge* (and hence, informational power) that can occur through technological discourse, this language should not be construed as unflinchingly determining individual knowledge.

Chapter 3 delves deeper into a specific contemporary genre of technology that a vast number of users have adopted: social media sites (SMSs). Chapter 3 first defines

social media sites, identifying some common technical characteristics of these sites, some of the common economics of these sites, and some practices of discourse creation shared by the various platforms' purveyors. After this introductory description, Chapter 3 surveys how various scholars have conceptualized the operation of power within the user-SMS relationship. After reviewing scholarly work that highlights both positive and negative outcomes for user power in this relationship, Chapter 3 observes that many of the negative outcomes highlighted by scholars are associated with particular economic practices of SMS purveyors (most notably, commodification of user-generated content), technical configurations of the platforms (in particular what Langlois et al. (2009) call "code politics"), and technological discourse that surround the technologies. In looking across a broad array of concerns regarding user power on SMSs, Chapter 3 argues that one of the problems scholars consistently identify fundamentally relates to users' (lack of) power in relation to the information flows on these platforms. However, the body of literature that identifies this issue appears to only tacitly recognize it as a problem of users' *informational* power. Further, the scholarly work that highlights concerns about users' power is largely piecemeal, rarely addressing power, user knowledge, and technological discourse in tandem. The work that is relevant in this area often lacks more empirical components such as detailed studies into users' states of knowledge regarding information flows on SMSs or analytic work that explores how information flows are depicted in the technological discourse that surrounds many of these platforms. From this, Chapter 3 argues that further study into the interconnection between user informational power (vis-à-vis *principles-knowledge* of information flows) on SMSs and the role of technological discourse in the user-SMS relationship is needed, and that the

social media platform Twitter is a timely space to investigate these issues. In conclusion, Chapter 3 sets up the primary research questions of this dissertation: In the user-Twitter relationship, what is the state of Twitter users' informational power in regards to the informational flows of the platform? What *principles-knowledge* of information flows do users have and how does the technological discourse surrounding the site created by Twitter's business purveyors describe information flows and potentially impact users' *principles-knowledge* (and hence informational power)?

However, before this dissertation can proceed with an empirical study into Twitter users' knowledge of information flows and into an exploration of how the technological discourse surrounding Twitter describes information flows on the platform, it first needs to establish a baseline of how information actually flows across Twitter. In order to provide this descriptive account of how information flows on Twitter, this chapter first lays out an operational definition of information and information flows that will guide the descriptive work. After defining the operationalized terms, this chapter moves to articulate how information flows across the sociotechnical system of Twitter. Unfortunately, uncovering and describing information flows on Twitter is no easy feat. These flows are comprised and governed not just by hardware and software that make up the material technology, but are also shaped by many "non-technical" components such as terms of service, privacy policies, and business practices. To breakdown the sociotechnical object that is Twitter and render visible the information flows of the platform in a methodical manner that takes into consideration both the technology and the social elements of the system, this chapter turns to the analytical framework put forth by van Dijck (2013) in her critical history of social media, *The Culture of Connectivity*.

While this framework does not provide an account of every single piece of the sociotechnical system that makes up a social media site, van Dijck's framework does serve as a practical toolkit for identifying and breaking apart some of the salient, yet interconnected components of social media platforms such as data, metadata, algorithms, protocols, defaults, informational content, users, business models, platform ownership, and governing practices. This analytical framework helps this chapter articulate how information flows through Twitter at multiple levels of abstraction in a technical sense, while at the same time helping to unpack the more social and political economic bases of information flows on the platform. Through van Dijck's framework, this chapter provides a descriptive account of Twitter as a sociotechnical system and traces the information flows present on Twitter.

Definitions

What is information?

Information is a complex term that evades easy definition. Across disciplines such as economics, political science, communications, information science, computer science, and physics (among others), there have been multiple, often competing, definitions for the term "information" (Braman, 1989; Buckland, 1991; Machlup, 1983). Even within a single discipline, there are often numerous ways that information is theorized and defined (Bates, 2005, 2006). Conceptualizations of information that focus on a single characteristic or phenomenon related to information can be problematic as they may inadvertently exclude other critical considerations, thereby drawing the boundaries around what is considered information too narrowly.

Braman (1989) offers an approach to conceptualizing information that quells some of the problems associated with singular definitions of information. In an article geared towards policymakers, Braman suggests a hierarchical taxonomic approach to information that includes conceptualizing information as a resource, as a commodity, as a perception of a pattern, and as a constitutive force in society. Each definition within the taxonomy has its own strengths and weaknesses, but importantly, Braman argues that the decision to use a particular part of the hierarchy should be driven by the contextual circumstances in which information is being understood, studied, and/or governed. She writes, "This hierarchy is based on differences in level of scope (how broad a range of social phenomena is incorporated into the concept) and complexity (how finely and variously articulated is the social organization that appears through the lens of the particular definition)" (1989, p. 235). By adopting a particular approach to information from within the taxonomy, different concerns and questions emerge in relation to the scope and complexity of the system information is being understood within.

The different approaches to information within Braman's taxonomy—when applied to Twitter—yield different concerns and questions regarding the information flows present on the platform. Taking a resource-based approach to defining information allows one to consider how particular pieces of information move throughout the Twitter ecosystem as a kind of good. Information flows could then be conceptualized as the conduits or pathways through which allocations of goods are made. The kinds of questions that emerge from such a view include questions such as "Where is information in this system and where is it going?" or "Who has access to information and information flows, and who does not?"

Approaching information as a commodity facilitates considering it as a kind of economic good or product. This view expands the scope of social phenomenon that is at play beyond that of just "resources." The added social elements of value and labor come more directly into focus. In defining information as a commodity rather than as just a resource, the questions that emerge expand to include questions such as: "What are the socio-economic factors involved with the allocation of these resources?" or "How do the economic factors involved in the production of this information shape how it flows?" When the commodity definition is applied to information flows, these flows can be better understood as part of the information production chain that helps bring the goods users produce (user-generated information) to a marketplace.

Viewing information as perception of a pattern yields a perspective that orients one towards considering how information from within Twitter can be used to reduce uncertainty. For example, information from Twitter can be used (to some degree) to poll public sentiment about a particular news story, brand, or product. Some stock market investors use Twitter data to help reduce their uncertainty about what direction a particular stock might go in. With this definition of information in mind, one might ask questions about how the information flows become part of efforts where uncertainty reduction is important, such as in decision making processes. This view also surfaces some of the implications of power for those who have access to Twitter's information flows versus those who do not.

Finally, defining information as a constitutive force in society allows one to consider how information on Twitter might have an active role in shaping societal contexts. This view of information facilitates considering not just the role of information

such as Tweets being used to help reduce uncertainty and to make decisions, but also how the presence of that information may shape or influence the kinds of questions that are asked in the first place. In application then, adopting this perspective may open up questions about how Tweets may contribute to structuring knowledge of the world. One might ask, for example, how knowledge of protest movements gained through Tweets impacts responses to these movements. Studying information flows with this definition in mind might open up questions such as: "How do the information flows of Twitter actively shape politics?"

Ultimately, as the concern of this dissertation is users' informational power vis-à-vis their *principles-knowledge* of information flows, the two definitions of information within Braman's taxonomy that are most applicable to this study are those of information as a resource and information as a commodity. Treating information on Twitter as resource facilitates the conceptualization of information flows as conduits through which a resource moves from point A to point B (or from A to B to C…). Recognizing information on Twitter additionally as a commodity allows this study to consider the broader political economic conditions in which this information is produced and distributed, and how users' knowledge of these flows fits into this production process. While much could be gained from research that considers information on Twitter as perception of a pattern or as a constitutive force in society, those are very different projects from this one. This dissertation's chief concern with users' knowledge of information flows at a *principles-knowledge* level and role of technological discourse in this process renders the other two parts of the hierarchical taxonomy less applicable for now.

What is an information flow?

Just as the term "information" has a variety of definitions that have emerged within various disciplinary approaches, "information flows" have a similar complicated history. Disciplines such as computer science, political science, economics, information science, communications, and media studies all inherently deal with information flows as embedded within a particular context, system, or medium. As the context for information flows in this dissertation have to do with a communication medium, and as this dissertation is concerned with information in its resource and commodity form, this project will explore information flows as they have been defined and incorporated in transmission models of communication.

The Shannon-Weaver model of communication (1949) is frequently touted as, "one of the main seeds out of which Communication Studies has grown" (Fiske, 1982, p. 6). It is a simple, linear communication model that incorporates information flows as they operate across the constituent parts of a sender, message, transmission, noise, channel, reception, and receiver. This is shown in Figure 2.

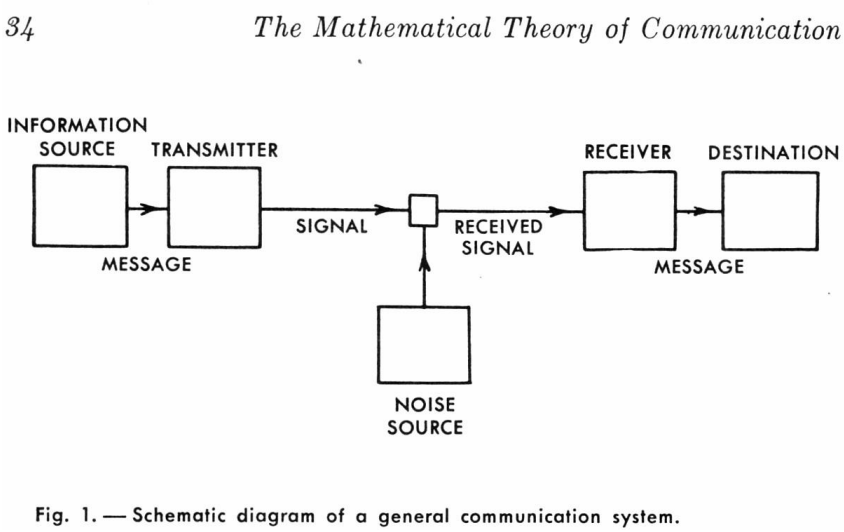


Figure 2. The Shannon-Weaver mathematical theory of communication. (Shannon & Weaver, 1949, p. 34).

Information flow occurs as a sender (information source) creates a message, transmits it through a transmitter across a channel which may be subject to noise, to a receiver which propagates the message to a destination. Within Shannon and Weaver's model, information is treated as both "perception of a pattern" (as the message can be subject to noise) and as a resource moving from sender to receiver through the conduit of a channel.

Despite it becoming one of the main "seeds" of communication studies, Shannon and Weaver's model has been critiqued as a poor general model of communication. It has been described as overly simplistic and linear, particularly because it does not take into account the active role of receivers in interpreting messages and lacks a consideration of feedback as part of the communication process (Chandler, 2011). Perhaps this is not an unreasonable critique as Shannon designed this model for mathematical and technical modeling[20] rather than attempting to create a general model of communication. Later communication models, such as Berlo's (1960) Sender-Message-Channel-Receiver

[20] For example, measuring uncertainty (entropy) in information within a communications system or measuring potential throughput of a channel.

model of communication and Barnlund's (1970) transactional model of communication built on Shannon and Weaver's work, taking into account different contextual factors in the communications process including sense-making and signification.

However, despite its criticisms, Shannon and Weaver's model makes for an excellent starting point for conceptualizing information flows within Twitter. This is because this dissertation is interested in information flows as they exist in the sense of linear conduits that move information as a resource/commodity from one point to another. More specifically, this dissertation is fundamentally interested in what *principles-knowledge* users have of these conduits, how that user knowledge is tied to technological discourse, and ultimately how a user's knowledge informs that user's informational power. As this dissertation is not treating information as a perception of a pattern, the only part of Shannon and Weaver's model that is not relevant to this study is noise. Therefore, for the purpose of this dissertation, an information flow is defined as: the means by which information, as a resource/commodity, is transmitted from a sender towards a receiver. The four critical elements of describing any information flow on Twitter, therefore, are: 1) the means/mechanism/channel by which information "flows," 2) the information, 3) the sender, and 4) the potential receivers. This definition is intentionally simplistic and does not take into account additional considerations such as sense-making, reciprocal communication, or how a receiver makes use of the information. This is diagrammed in Figure 3.

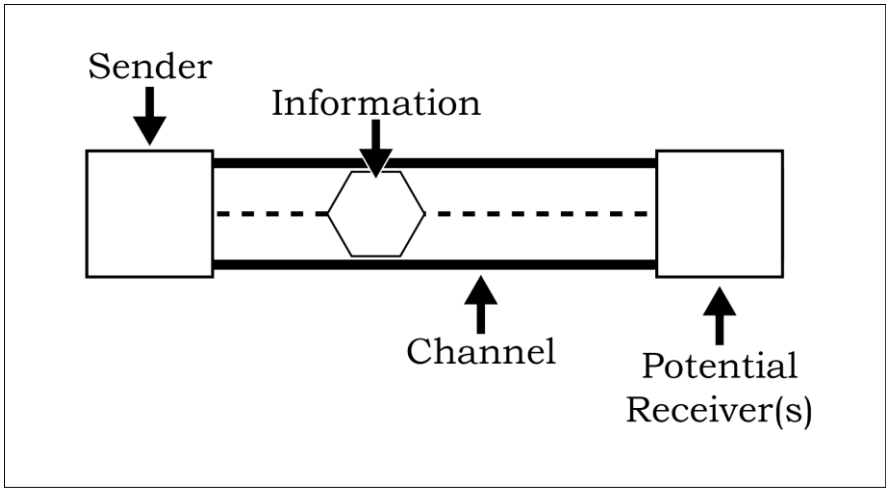


Figure 3. Diagram of elements of an information flow.

It is important to note that the transmission of information on Twitter is a multi-step process. Twitter is *not* being conceptualized here as merely the channel between User A and User B (as sender and receiver, or follower/followee). Instead, when User A creates a Tweet, that user must first communicate that Tweet to Twitter. Twitter, in this first step, is acting as receiver. Subsequently, as a second step, Twitter then acts as sender, making that information available to many other potential receivers. Further, overt intentionality should not inherently be ascribed to the transmission of information from a sender to receiver. As this chapter later describes, there is much more to a "Tweet" than just the 140 characters of text that a user may intentionally create. There is often associated metadata that a Tweet creator (acting as sender) may transmit that they may be totally unaware that they are generating.

Deconstructing Twitter

Twitter is a complex technology made up of not just material objects such as servers and hardware, but also more intangible components such as software, protocols, algorithms, terms of service, data, users, owners, governing documents, business practices, etc. While information flows on Twitter are only comprised of four constitutive

parts (sender, information, channel, and receiver), multiple elements within the sociotechnical object shape the exact arrangement and configuration of information flows. In order to create a baseline account of how these different components constitute information flows, this study requires an analytical method for deconstructing Twitter.

In van Dijck's (2013) critical history of social media, she states the goal of her book is in "understanding the coevolution of social media platforms and sociality in the context of a rising culture of connectivity" (p. 28). However, understanding the co-evolution of these elements requires tracing these systems as complex sociotechnical objects. In order to illustrate the ongoing co-evolution between users and social media platforms, van Dijck suggests combining an actor-network theoretical approach along with a political economic approach in order to analyze the "dynamic intricacies of platforms" (p. 28). Even though the work of this dissertation is not in understanding the co-evolution of platforms and sociality, the analytical mechanisms van Dijck employs for breaking down sociotechnical objects are well suited for the task of explicating Twitter and the information flows of the platform. This is because, fundamentally, she is interested in exploring the same relationship between users and social platforms that this dissertation is, but rather than exploring how the user/SMS configuration gives rise to greater sociality, this dissertation is interested in users' informational power.

In arguing how one can deconstruct a social media platform, van Dijck states that it is important to consider two different elements of social media: the "techno-cultural constructs and socioeconomic structures" (p. 28). These two are separable elements, but also influence and inform each other. She further breaks down each of these elements into three constituent parts. In order to evaluate the "techno-cultural constructs," she

focuses on three elements, the functioning of the technology itself, the users, and the informational content of a platform. In order to explicate the socioeconomic structures, van Dijck focuses on the ownership status of the platform, its governance, and the business models that undergird the technology. It is through articulating all six elements and tracing their interrelations that van Dijck makes her case regarding the coevolution of social media platforms and sociality.

Applying van Dijck's framework to the description of information flows on Twitter appears relatively straightforward. By accounting for the "techno-cultural" elements of Twitter, this study can inherently give a descriptive account of the information flows on the platform in their current form. For example, the receivers and senders of content within the circuit of an information flow would be what van Dijck considers as "users." The technical infrastructure can be considered as the "channel" of an information flow. Finally, the content that van Dijck is concerned with can be considered as the particular kinds of information present in the information flows. This application of van Dijck's terminology to the model of information flows can be seen in Figure 4.

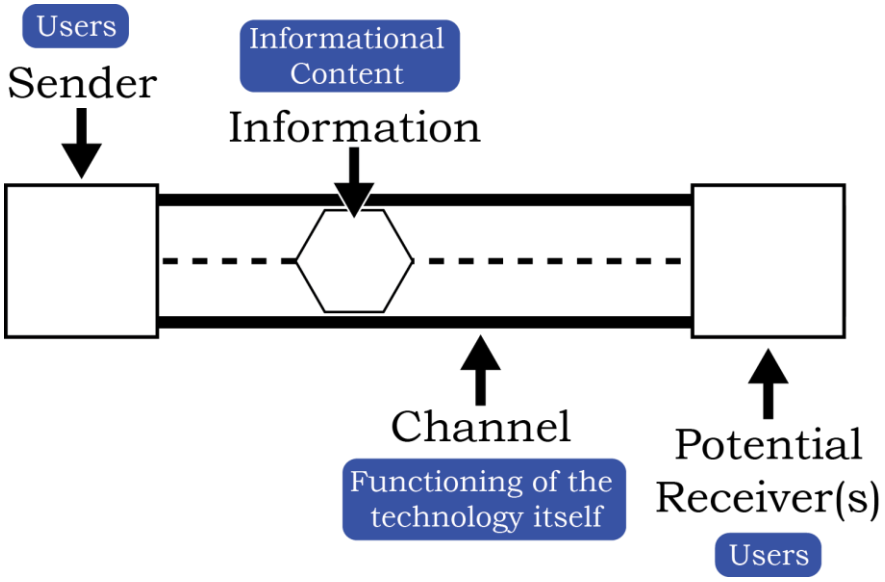


Figure 4. Diagram of elements of an information flow with van Dijck's terminology applied.

However, in order to understand why the information flows on Twitter exist in the way that they do, it is also necessary to examine the socioeconomic structures of Twitter, focusing on its ownership status, governance, and the business models of the platform. These pieces are not pictured in Figure 4, but greatly influence the exact arrangement of users, informational content, and the functioning of the technology as they exist on Twitter. Each of these six areas is explored in greater detail next.

The techno-cultural dimension: Technology.

The role of material technology is perhaps the most obvious element of information flows on the Twitter platform. It provides the mechanism through which information takes its particular shape on Twitter and serves as the material structure that governs how information gets from point A to point B. However, the "technology itself" is actually made up of many different components and must be broken down further in order to be described coherently. Van Dijck relies on a conceptual framework for exploring the technological side of social media platforms that breaks it down into five

interrelated elements: data/metadata, algorithms, protocols, interfaces, and defaults (p. 30).

According to van Dijck, data can be understood as a type of signal in a form suitable for use within a given system. Social media platforms often have extensive formatting rules and code that configures the shape of data within its system. For example, Twitter famously limits the length of Tweets to 140-characters. Van Dijck notes that often related to, and associated with, data is metadata. Metadata is structured data used to describe, explain, and locate other data. Metadata are an important part of social media platforms because users often rely on it to discover or manage other data. For example, Twitter's search algorithm relies on both data and metadata to help searchers find Tweets. For the purposes of this dissertation, the way data and metadata are structured on Twitter is important because these structures essentially create an ontology for the kinds and types of information that exist within the platforms and within information flows.

Social media platforms, van Dijck notes, are not just dumb terminals through which data and metadata flow. These platforms also exercise "computational power," manipulating data and generating new data through the use of algorithms. An algorithm, essentially, is a set of instructions or code that produces a certain output based on given inputs. Examples of algorithms include things such as Facebook's "People You May Know" feature and Twitter's "Trending Topics" feature. Each of these algorithms takes a set of data from within the system as an input and then produces a set of outputs (which in these cases are made visible to users through the social media's respective interfaces). One of the difficulties of describing algorithms is that their exact inner workings are often

not publicly accessible. For example, while one can observe the outputs of Facebook's "People You May Know" algorithm, the exact inner-workings of the code that generates the recommendations remain inaccessible to users. Algorithms are an important component of information flows, as they have an additive effect on the total amount of information contained within the system. For example, when an individual sends a "Tweet" through Twitter, Twitter may do computational processing on this Tweet, creating through its algorithms new information that is added to the initial message or becomes accessible through the interface.

In addition to algorithms, social media platforms often rely on protocols as part of their technical structure. Based on the work of Galloway (2006), van Dijck writes that protocols are technical sets of rules (or "scripts") that govern user behavior within the logic of a platform. For example, Facebook provides many scripts that guide user action, such as the code that allows a user "like" something, but does not allow for a user to "dislike" something. Because of the scripted nature of protocols, individual users will often find difficulty in engaging in behavior with the platform outside of these prescribed scripts. Van Dijck suggests that protocol can "impose a hegemonic logic onto a mediated social practice" (p. 31).[21] What is important about protocols for the purposes of this dissertation however, is how protocols may act as scripts that help regulate and shape the form of the information that flows through Twitter.

Interfaces are often closely tied to algorithms and protocols, as they are the objects that serve to link software, hardware, and users to data (Fuller, 2008). Interfaces often contain the elements that steer users towards particular protocol governed behavior.

[21] This is not to say that protocol is entirely deterministic in nature however, as "protocological control by platform owners often meets protocological resistance from defiant users" (van Dijck, 2013, p. 31).

As van Dijck notes, "Interfaces…are an area of control where the meaning of coded information gets translated into directives for specific user actions" (p. 31). In addition to website based interfaces, van Dijck notes that application program interfaces (APIs), such as those found on Facebook and Twitter, are a key kind of interface for social media platforms. Many SMSs provide different kinds of interfaces for different kinds of users. For the purposes of this dissertation, interfaces are important as they are the means of user access to information flows within the social media platform of Twitter.

Finally, echoing the comments of scholars such as Lewis, Kaufman and Christakis (2008) and Tufekci (2008), van Dijck notes the important role that interface defaults play for channeling user behavior and information creation in specific ways. For example, she notes that Facebook's decision to make all information posted by a user public by default contains an inherent ideological maneuver, and that "if changing a default takes effort, users are more likely to conform to the site's decision architecture" (p. 32). This is an important observation as van Dijck notes "algorithms, protocols, and defaults profoundly shape the cultural experiences of people active on social media platforms" (p. 32). In the context of Twitter, defaults play an important part in shaping what information flows from which senders to which receivers.

A diagram of this dissertation's model of information flows updated with van Dijck's five-piece breakdown of "technology" concepts is provided in Figure 5.

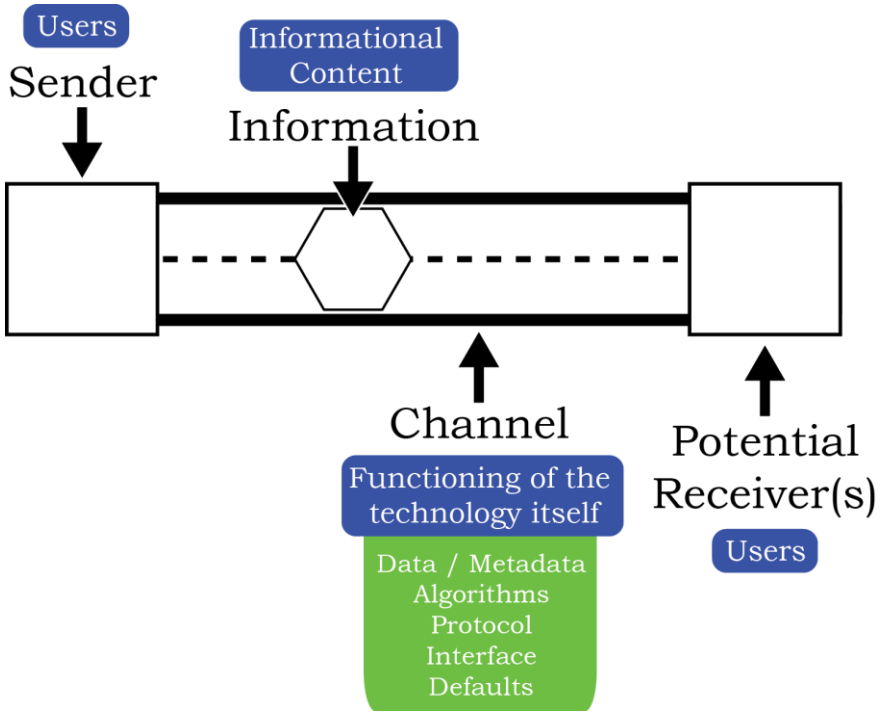


Figure 5. Diagram of elements of an information flow with components of channel from van Dijck's framework.

With the five different aspects of van Dijck's framework for analyzing technology now explained and, having given a brief introduction on their relevance to describing information flows on Twitter, this chapter next dives into the specific details of how these aspects manifest on Twitter. The next five subsections of this chapter break apart the technology of Twitter, describing it by its constituent parts of interface, protocol, algorithms, data/metadata, and defaults. The descriptions of these elements comes from a combination of the descriptions of the service given in Twitter's Terms of Service and associated policies, Twitter's technical documentation meant for application developers, from Twitter's developer blog, from the author of this dissertation's own examination of the technology, and from secondary sources.

Interface.

Since its creation in 2006, the interfaces that Twitter offers have undergone a constant evolution. While the look and feel of these interfaces has changed over time to support new additional functions, the core principle of Twitter as a micro-blogging platform has stayed consistent. Today, Twitter offers multiple ways of interacting with the Twitter ecosystem, including through its web-based interface, through its application layer interfaces (APIs), and through buttons and widgets that are embeddable in third-party websites. Of registered Twitter users, roughly one-third rely on the web-interface, whereas two-thirds rely on applications built on-top of the API frameworks to interact with the Twitter ecosystem, such as mobile apps and desktop applications (Beevolve, 2012).

Web-interface.

The web-based interface of Twitter.com provides a way for registered users to both consume the stream of 140-character messages that others are posting and to create and share their own content. Nonregistered users can still read most content posted to Twitter.com and can use the search tools, but cannot share their own messages through the service, and do not have the same abilities to access portions of the web-based interface as registered users. The differences among registered and nonregistered users and their respective abilities to participate in information flows on Twitter are discussed further in the subsection on users.

The Twitter.com interface that registered users interact with has changed significantly since its original design in 2006. In the current (April, 2014) design, at the top left of the website, there are four main "tabs" that a registered user can interact with: the "Home" tab, the "Notifications" tab, the "#Discover" tab, and the "Me" tab.

Regardless of what "tab" a user is in, there are a set of controls on the top right that will appear consistently: the "Search Bar" and three buttons that allow a user to access what are called "Direct Messages," a "Settings" page, and a button that allows them to compose a new Tweet. These items can all be seen labeled at the top of Figure 6. For the purposes of this discussion, this analysis will focus only on the "Home" tab. Figure 6 displays the "Home" screen tab which can more or less be considered the main-interface for registered users engaging the Twitter.com website.

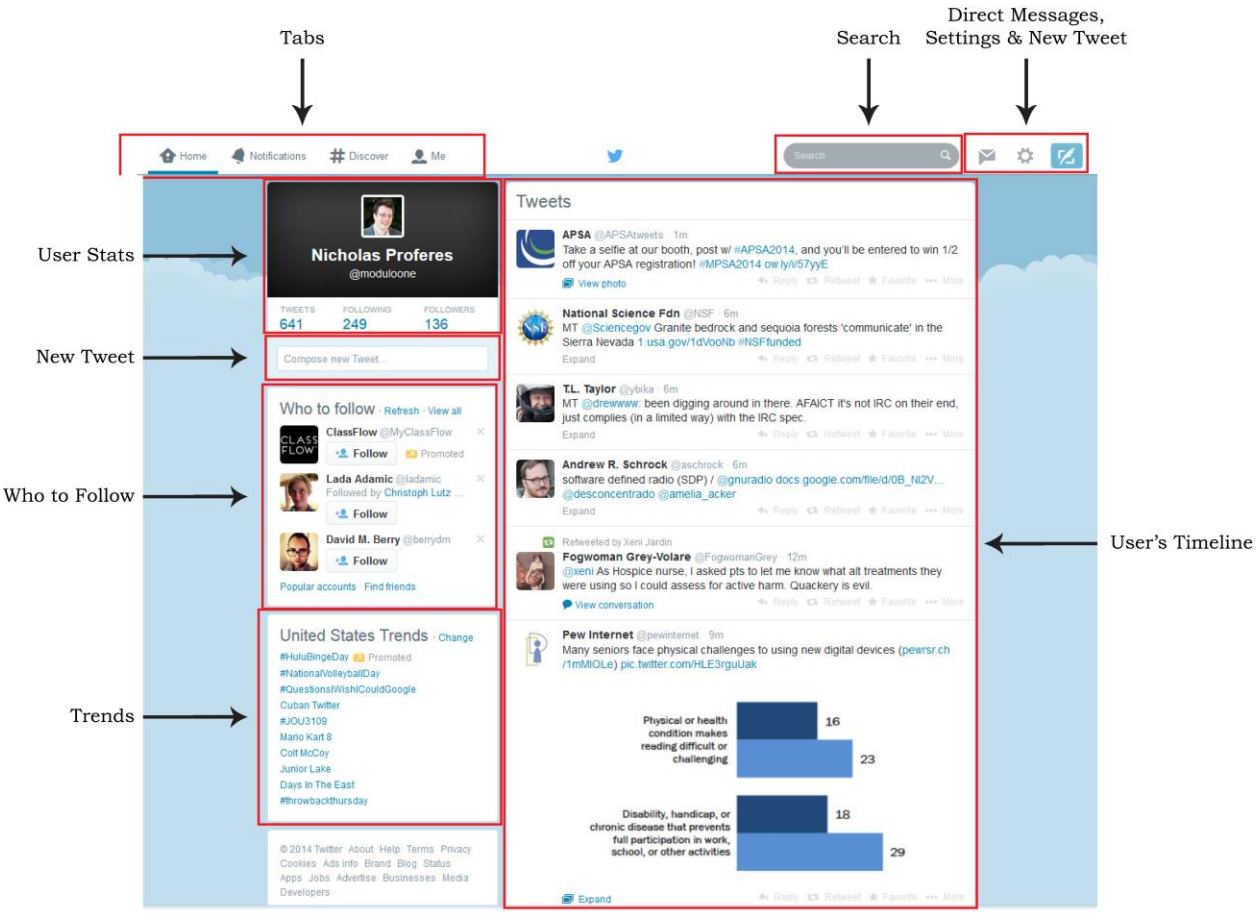


Figure 6. The Twitter.com web-interface "Home" tab.

As can be seen from Figure 6, there are five main areas that users can interact with in some capacity on the Home tab: the "User Stats" area, the "New Tweet" area, the "Who

to Follow" area, the "Trends" area, and the "User's Timeline." The "User Stats" area provides basic information about the user that is logged into the Twitter system, the number of Tweets that user has generated, the number of people that user follows, and the number of people who are following that user. Below this area is a text-box that allows a user to generate a new Tweet. Below that, the interface provides a list of suggested users to follow (generated by one of Twitter's internal algorithms) as well as a "Find Friends" tool that can be used to discover other users by importing e-mail contact lists. Below that, Twitter provides information on "Trending Topics," which are determined by an algorithm. Finally, to the right, the user timeline contains a reverse chronological stream of messages from that user and from the users that individual "follows" with the most recently created messages appearing at the top.

Historically, the question that historically appeared near or in the "New Tweet" area was "What are you doing right now?" This question was eventually replaced by, "What's happening?" and, as of the time of writing, has been replaced with the much more simple and less inquisitive "Compose new Tweet…" When a user clicks into this text box, additional functions appear that allow users to upload a photo, allow users to include a location, allow users to see the number of characters he or she has left for the Tweet (as there is a 140-character limit on messages within Twitter), and allows users to "Tweet" the message, thereby transmitting the uploaded text, photo and/or location information into Twitter. This is shown in Figure 7.

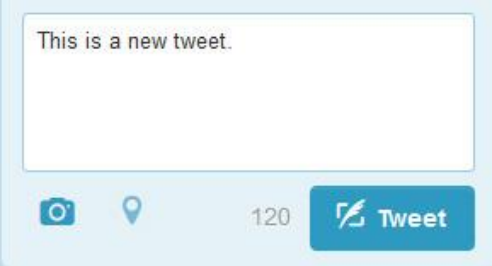


Figure 7. Panel to compose a new Tweet.

Once a message has been created and the user has pressed "Tweet," the message is transmitted to Twitter. Once in Twitter's hands, a number of things happen. First, Twitter's algorithms go to work on the message, parsing it for a number of different purposes. One particularly notable use (as it relates to another area of the Home-tab) is that Twitter parses the text of the Tweet in order to algorithmically determine what "Trending Topics" should appear in the "trends" area. Trending topics are popular discussion points present within Tweets found in different geographic areas (such as "Trends in Milwaukee," "Trends in Wisconsin," "Trends in the United States," and "Global Trends"). Twitter describes trending topics this way:

> Trends are determined by an algorithm and, by default, are tailored for you based on who you follow and your location. This algorithm identifies topics that are immediately popular, rather than topics that have been popular for a while or on a daily basis, to help you discover the hottest emerging topics of discussion on Twitter that matter most to you. (twitter.com, 2014s)

The trending topics algorithm is described in greater detail in the subsection on algorithms. When a user clicks on a trending topic in the "trends" area, they are taken to a search page that displays a stream of the most popular and most recent Tweets that contain mention of the particular trending term within a given geographic area.

Once a Tweet has been sent to Twitter, Twitter then makes that message available within the author's Twitter profile page (which may be either public or protected), within the author's own timeline, and within the timelines of users who follow that author. It is

also made available through Twitter's APIs, through which applications and developers gain access to Tweets (the APIs will be discussed momentarily). Each Tweet that a user generates is given a unique identifier (discussed further in the data section) as well as its own unique URL at which it can be accessed. How a Tweet appears within the unique URL is depicted in Figure 8.

Figure 8. An individual Tweet with a unique URL of https://twitter.com/moduloone/status/431515847224332288

It is important to note however, that there are some limits to the information that can be accessed through the web-interface. Perhaps most notably, there is a limit to the number of Tweets that can be loaded into a user's timeline: 3200. This means that a user cannot go back in time into the Tweet histories of other users beyond a certain point. It also means that the user cannot, through the timeline, access their own older messages beyond a certain point, unless the user has stored the Tweet URL or Tweet ID. In practice then, if a user has created more than 3200 messages, they will be unable to access the 3201st message through the timeline interface unless they know the Tweet URL or the Tweet ID. In 2013, Twitter did introduce a "download personal archive" feature that is

part of the "user settings page." The download personal archive feature does allow for the bulk retrieval of older messages that an individual has authored, but this data is delivered in a compressed format over e-mail, and not within the web-interface itself. Further, as will be described in more detail in the discussion of the APIs and in the section on data on Twitter, some of the metadata that Twitter maintains regarding Tweets is formatted differently in the APIs than in the web-interface.

Application programming interfaces.

Twitter's "Application Programming Interfaces" (APIs) are a set of interfaces that Twitter offers that allows programmers, developers, and applications to interact with Twitter's services and data. Twitter describes its APIs this way:

> An API is a defined way for a program to accomplish a task, usually by retrieving or modifying data. In Twitter's case, we provide an API method for just about every feature you can see on our website. Programmers use the Twitter API to make applications, websites, widgets, and other projects that interact with Twitter. Programs talk to the Twitter API over HTTP, the same protocol that your browser uses to visit and interact with web pages. (twitter.com, 2014g)

Twitter writes in its "Developer Rules of the Road" that it provides this interface to maintain "an open platform that supports the millions of people around the world who are sharing and discovering what's happening now. We want to empower our ecosystem partners to build valuable businesses around the information flowing through Twitter" (twitter.com, 2013a). Most stand-alone applications that interact with Twitter, such as TweetDeck and HootSuite, use the Twitter APIs to get data to and from Twitter. Many API users also use interfaces as a bulk data collection and retrieval tool. For example, many researchers rely on the APIs to collect data as part of their studies (Gaffney & Puschmann, 2013; Zimmer & Proferes, 2014).

Historically, Twitter has offered three separate "types" of APIs: the REST APIs, the Search APIs, and the Streaming APIs. In 2013, Twitter began to unify these three separate APIs into a single interface, API Version 1.1, and scheduled the discontinuation of the older APIs for June of 2014. However, API Version 1.1 still contains most of the same functionality as the older interfaces, just with some modifications to the rate at which data can be pulled and posted to Twitter and a move away from delivering data in Extensible Markup Language (XML) format in favor of the JavaScript Object Notation (JSON) format (twitter.com, 2013c). For the purposes of this dissertation, this chapter will briefly describe each of the three APIs individually, although with the acknowledgement that they are now being unified into a single framework.

In describing the APIs, Twitter summarizes what each API does this way:

> The Twitter REST API methods allow developers to access core Twitter data. This includes update timelines, status data, and user information. The Search API methods give developers methods to interact with Twitter Search and trends data. The concern for developers given this separation is the effects on rate limiting and output format. The Streaming API provides near real-time high-volume access to Tweets in sampled and filtered form. (twitter.com, 2012)

Each respective API is important because each one offers a distinct set of functionalities and characteristics, allows different levels of access to information on Twitter, and has different costs and use rights. As a result, each API plays a distinct role within the information flows on Twitter.

The REST API serves as an interface that allows authorized users, through the use of "methods," to request or post a limited set of data to the Twitter ecosystem. A full accounting of the hundred-plus information posting and retrieval methods offered in API Version 1.1 through the REST API is listed in Appendix 1. The REST API is a public API in the sense that any registered user can request access to it for free. However, in

order to make use of it, users must have some level of familiarity with programming or have pre-built software that can interact with the APIs in order to be able to make efficient use of the interface. In the older versions of the REST API, users could make data request calls to API without having to be registered with Twitter. In the new version of the API, this is no longer the case.

The REST API offers a wide variety of methods for interacting with data in Twitter; however, many of these methods of data retrieval are not as easily achieved through other interfaces, such as the web-interface. For example, the "GET friends/ids" method (detailed in Appendix 1) allows an API user to retrieve a bulk list of all the user IDs of the authenticated users' friends. While it is possible to look this information up manually through the web-interface, there is no automated tool to do it. API users, however, are faced with certain restrictions on their use of the interface, for example, being limited in the number and frequency of calls they can make to the REST APIs. Depending on the method invoked, users of the REST API are limited to somewhere between 15 calls per 15 minute-window and 180 calls per 15 minute-window (twitter.com, 2014i). What is important to observe about this limiting is that for applications that are attempting to gather large amounts of data or are attempting to interact with data on Twitter in "real-time," the REST APIs are a less than ideal interface. Information flows through this part of the channel are essentially rate-delayed. Twitter's own documentation of the APIs notes that if a programmer needs more "real-time" interaction with the Twitter data ecosystem, they should instead use the Streaming API.

The second type of API that Twitter offers is called the Streaming API. Twitter describes the Streaming API by noting that it "provides low-latency high-volume access

to Tweets" (twitter.com, 2014l). Because it offers higher volume access to Tweets and is not rate-limited in the same way as the REST APIs, many researchers and data-miners gain access to data within the Twitter ecosystem through the Streaming API. In writing about the Streaming API, Gaffney and Puschmann (2013) observe that unlike traditional APIs which require a "pull" request from the user, the Streaming API operates on a "push" basis, where "data is constantly flowing from the requested URL (the endpoint), and it is up to the [user] to develop or employ tools that maintain a persistent connection to this stream of data while simultaneously processing it" (Gaffney & Puschmann, 2013, p. 56). Within the Streaming API, there are several specific data-streams that Twitter offers access to. These are described in Table 1.

Table 1

Streaming API Data Streams, Based on Description Given by Twitter, Inc. (twitter.com, 2014k)

Stream	Description
Public streams	Streams of the public data flowing through Twitter. Suitable for following specific users or topics, and data mining.
User streams	Single-user streams, containing roughly all of the data corresponding with a single user's view of Twitter.
Site streams	The multi-user version of user streams. Site streams are intended for servers which must connect to Twitter on behalf of many users.

Notably, the Streaming APIs are delivered in three "bandwidths": "spritzer," "garden-hose," and "firehose" which deliver 1%, 10% and 100% of all Tweets posted to the Twitter ecosystem (Gaffney & Puschmann, 2013). All registered API users are automatically granted rights to the "spritzer" channel; however, access to the "garden-hose" or "firehose" requires an additional agreement with Twitter. These agreements

often come in the form of a business relationship. Firehose access has historically been particularly sought after by third-party developers. In a developer forum on Twitter, one of Twitter's platform product managers, Taylor Singletary, writes in response to how to get access to the "firehose":

> Firehose access is very hard to come by and potentially very expensive to realistically consume. Many businesses that gain access to the Firehose do so through an evolutionary set of steps, beginning by working with the most basic levels of the streaming API (1% of the firehose), validating their product, business model, and value to the Twitter ecosystem before working their way up the various access levels. It also depends on the type of product you're building. Developers want streaming access for different reasons. Some of those reasons require going through one of two resellers of Twitter firehose data, @Gnip or @DataSift. (Singletary, 2012)

As a result of the stratification of data access, the garden-hose and firehose are considered to be less publicly accessible. boyd and Crawford (2011) observe that, as a result of the gradated access, Twitter essentially only makes a small percentage of the data it has available through the public APIs. Essentially, the pool of potential receivers for this part of the channel is shaped by the business relationships that have been struck between Twitter and those third-parties. The exact costs of firehose access today are not well known; however, Small et al. (2012) noted that Google reportedly paid $15 million dollars to access the full stream of all public Tweets in 2010 for just that year.

Third and finally, there is the Search API. The Search API was originally the only method for searching the public stream of Tweets for particular mentions of hashtags or terms through the APIs. As the APIs have been reworked, this is no longer the case. Historically, search on Twitter was actually provided by a third-party, Summize, Inc., and not by Twitter (twitter.com, 2012). In 2008, Twitter purchased Summize; however, Twitter had difficulty fully integrating the search API into the codebase of the site. As a

result, the Search API has historically been separate from the REST API, despite the fact that they are closely related and often used in tandem. As part of the larger, reworked API Version 1.1, search is now integrated into the REST API. In describing the Search API, Twitter notes:

> It allows queries against the indices of recent or popular Tweets and behaves similarily [sic] to, but not exactly like the Search feature available in Twitter mobile or web clients, such as Twitter.com search. Before getting involved, it's important to know that the Search API is focused on relevance and not completeness. This means that some Tweets and users may be missing from search results. If you want to match for completeness you should consider using a Streaming API instead. (twitter.com, 2013f)

The Search APIs have been noted as particularly troublesome for researchers and data-miners as, "Data loosely falls off of the search system within a week of being posted, and no reliable information is available on completeness" (Gaffney & Puschmann, 2013, p. 60). However, the Search API does offer a unique way to access particular sets of data from within the Twitter ecosystem. The Search APIs allows for requests of public Tweets based on certain kinds of metadata contained within the Tweet (see the section on data on Twitter for an explanation of the different kinds of metadata on Twitter). For example, searches can specify that they are only interested in Tweets written in specific languages (as an identification flag is given in the Tweet metadata) and searches can specify that they are interested in Tweets generated within specific geographic areas (geolocation metadata). When searching by geolocation, a searcher provides a latitude, longitude and radius area, and then, "the search API will first attempt to find Tweets which have lat/long within the queried geocode, and in case of not having success, it will attempt to find Tweets created by users whose profile location can be reverse geocoded into a lat/long within the queried geocode" (twitter.com, 2013f). This is

notable because even if a user does not choose to include a geolocation as part of a Tweet, their Tweets may still be returned in geolocation-searches based on the location information they provided in their profile.

This introduction to the APIs concludes with a small number of observations about the larger data ecosystem created through these services. First, the moment that a user posts to the Twitter ecosystem, either by engaging the web-based interface or by posting information to Twitter through the REST API, that data becomes "live" within Twitter's ecosystem and is made available through all three APIs. This means that, for receivers connected to the Streaming API, they receive data quickly after it has been generated. However, if a user goes back to modify or delete data on Twitter, Twitter modifies it within its own databases, but it cannot control those who have already collected or cached that information. Twitter does send notifications to third-parties when a Tweet has been deleted, but the third-parties may still be able to maintain the deleted Tweet in their own databases, though they risk running afoul of Twitter's policies in doing so. Further, as mentioned in Chapter 3 of this dissertation, in their critical evaluation of Twitter's APIs versus the standard web-based user interface, Puschmann and Burgess (2013) argue that users on Twitter lack true control over the information they create if they do not use the APIs. They conclude:

> In the current state, the ability of individual users to effectively interact with "their" Twitter data hinges on their ability to use the API, and on their understanding of its technical constraints. Beyond the technical know-how that is required to interact with the API, issues of scale arise: the Streaming API's approach to broadcasting data as it is posted to Twitter requires a very robust infrastructure as an endpoint for capturing information ... It follows that only corporate actors and regulators— who possess both the intellectual and financial resources to succeed in this race—can afford to participate, and that the emerging data market will be shaped according to their interests. End-users (both private

> individuals and non-profit institutions) are without a place in it, except in the role of passive producers of data. (p. 11)

boyd and Crawford (2012) make a similar observation, arguing that the stratification of the Streaming APIs essentially creates divisions among those who use the different APIs.

Buttons and widgets.

Buttons and widgets are tools that third-party websites can use to embed pieces of the Twitter interface into their own websites. These allow individuals browsing the third-party websites to more easily interact with elements of the Twitter ecosystem while simultaneously being located at the third-party site. Examples of the four kinds of buttons that are embeddable are shown in Figure 9.

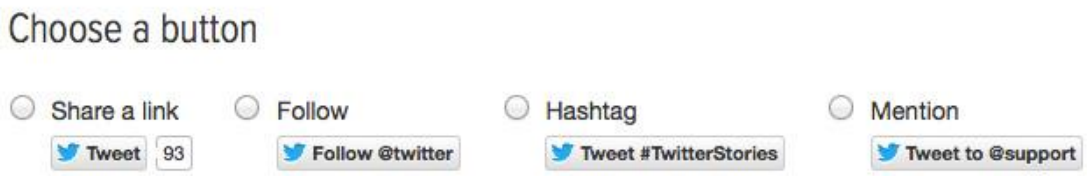


Figure 9. Twitter's embeddable buttons as found in (twitter.com, 2014n).

Many third-parties use these tools to help promote their own content, Twitter accounts, or specific hashtags within the ecosystem of Twitter. For example, the Tweet button (the button furthest left in Figure 9), "allows users to easily share your website with their followers" (twitter.com, 2013d). Similarly, widgets allow a registered Twitter user to embed particular elements of their own Twitter timelines within an external website they control. This allows registered Twitter users, for example, to embed a timeline of their own Tweets/retweets, their favorite Tweets, their lists, of particular searches or hashtags into their own websites. Twitter describes these widgets by stating:

> Embeddable timelines make it easy to syndicate any public Twitter timeline to your website with one line of code. Just like timelines on twitter.com, embeddable

> timelines are interactive and enable your visitors to reply, Retweet, and favorite Tweets directly from your pages" (twitter.com, 2014d).

Through both buttons and widgets, third-party websites can essentially become a branch-extension of the Twitter platform, facilitating the transfer of data to Twitter through these embeddable interfaces.

These buttons and widgets represent an important third way that data can become part of the Twitter ecosystem (twitter.com, 2013e). However, the data that Twitter receives from these sites is not necessarily just the data that users consciously upload through use of the buttons and widgets. Twitter indicates:

> Like many companies, Twitter receives log data from visits to websites that use our social widgets, such as our Tweet and Follow buttons. This log data may include information such as IP address, browser type, the referring web page, pages visited, cookies, and other interactions with the buttons or widgets as outlined in our privacy policy. (twitter.com, 2014g)

What is important about this statement is that the data that Twitter is able to collect from these widgets and buttons includes the URL of the website that a user is visiting. This happens regardless of whether or not the user actually interacts with the buttons. In writing about how the buttons and widgets interface with the larger Twitter data ecosystem, Harkinson (2013) writes:

> Much of the data Twitter collects about you doesn't actually come from Twitter. Consider the little "tweet" buttons embedded on websites all over the net. Those can also function as tracking devices. Any website with a "tweet" button—from Mother Jones to Playboy—automatically informs Twitter that you've arrived. Last year, Twitter announced that it would start using its knowledge of your internet browsing habits to better recommend people to follow on Twitter. (para. 3)

As Curtis (2012) observes, "Basically, every time you visit a site that has a follow button, a 'tweet this' button, or a hovercard, Twitter is recording your behavior. It is transparently watching your movements and storing them somewhere for later use" (para.

3). In a follow-up e-mail sent by Twitter's support team to Curtis, made public by Curtis, Twitter's representatives noted:

> To protect your privacy, we do not maintain browsing history. We start the process of deleting your visits to pages in the Twitter ecosystem after a maximum of 10 days. We only keep tailored suggestions for you, as explained in our privacy policy. (Curtis, 2012, para. 13)

What this highlights, however, is that information flows that involve data moving from a user to Twitter can occur not just when a user is consciously engaging the Twitter.com web-interface or the APIs, but may also occur while the user is out on the web, browsing third-party sites that happen to have these pieces of Twitter's channel embedded in them. This means that in conceptualizing information flows within Twitter, it is important to keep in mind that the "sender" in some cases may be an individual user who may be unaware that they are transmitting data to Twitter and may not consciously be interacting with the Twitter platform.

Defaults.

While social media platforms such as Facebook have received significant attention for their configurations of user default settings, Twitter has received less attention respectively, but has also made a number of similar and important decisions about the default settings of registered users. These default settings are a critical way that the information flows on Twitter are shaped, as research has indicated that many users never change the default settings that are chosen for them (boyd & Hargittai, 2010). This section will proceed by highlighting the default settings chosen for new users, registering through the web-based Twitter.com interface in early April of 2014.

When signing up for an account on Twitter, even before the account has been created, an individual has the opportunity to make choices about how their information

will flow to Twitter, with certain defaults already having been selected. For example, in the sign up process, by default, Twitter suggests keeping login credentials stored on the users' computer. Further, by default, "Tailor Twitter based on my recent website visits" is turned on.[22] In describing the "Tailor Twitter" feature, Twitter states:

> We determine the people you might enjoy following based on your recent visits to websites in the Twitter ecosystem (sites that have integrated Twitter buttons or widgets). Specifically, our feature works by suggesting people who are frequently followed by other Twitter users that visit the same websites. (twitter.com, 2014p)

The defaults as part of the sign-up process are shown in Figure 10.

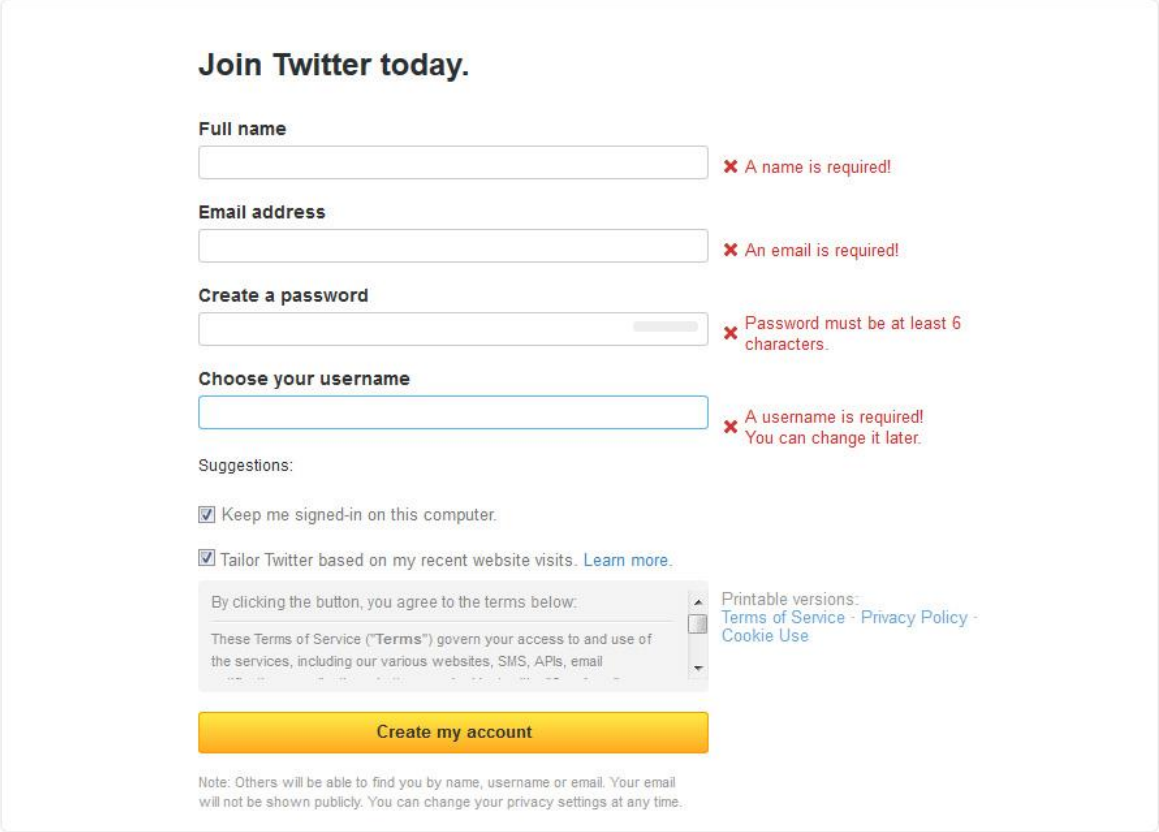


Figure 10. The Twitter.com new user registration page with "Suggestions" as defaults.

[22] It should be noted however, that this feature is actually not turned on by default if a user has the "Do Not Track" setting activated within their web-browser.

Once a user has completed the registration process (no other choices or defaults are present in the registration process), there are a number of automatically assigned default settings that are accessible through the user's "Settings" page. However, it is important to note that a user does not have to even ever look at their settings page before they can begin using the service. The only defaults that a user is confronted with before they are able to use the service are those that are part of the registration page. On the user's "Settings" page (which can be seen on the left hand side of Figure 11) there are nine categories of settings that a user can interact with: "Account," "Security and privacy," "Password," "Mobile," "Email notifications," "Profile," "Design," "Apps," and "Widgets." This section will briefly highlight the defaults selections made for users on each page, where applicable.

On the "Account" section (shown on the right hand side of Figure 11), there are two settings related to each other that both have default selections made for the new user. These have to do with what Twitter identifies as "sensitive media." The first setting, which is turned off by default, states, "Do not inform me before showing media that may be sensitive." By default, users are given a warning if media that others have uploaded (pictures, links, or movies) contains content that has been identified as sensitive by the uploader or by other users. Related is a second setting that states, "Mark media I tweet as containing material that may be sensitive." By default, this setting is turned off, meaning that user uploaded content is not marked as sensitive by default. In explaining why a user might want to turn this setting to the "On" position in the settings explanation page, Twitter states: "If you upload Tweet media that might be considered sensitive content

such as nudity, violence, or medical procedures, you should consider applying the account setting ‘Mark my media as containing sensitive content’” (twitter.com, 2014v).

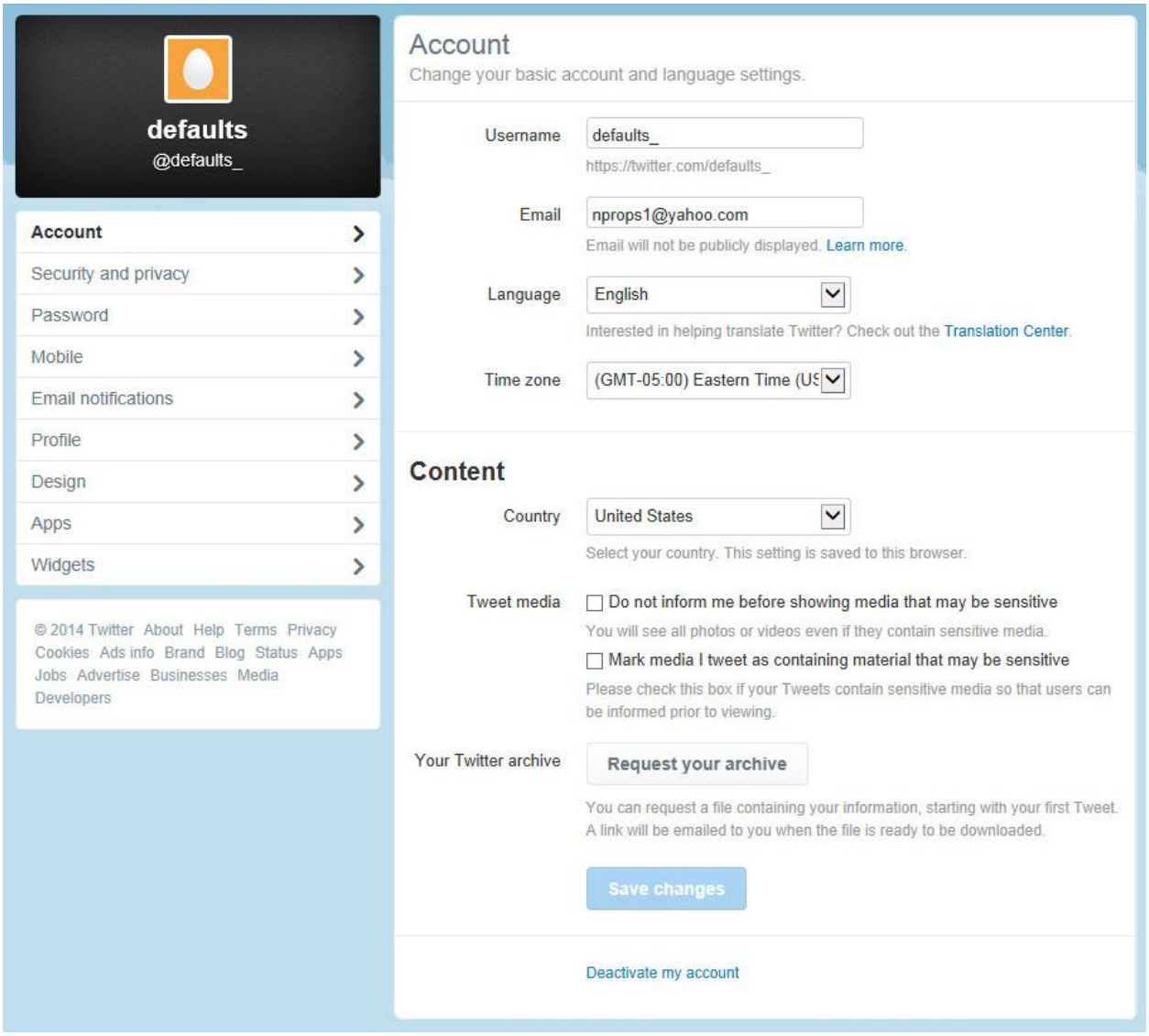


Figure 11. The user “Account” settings page.

Under the “Security and privacy” tab, there are a number of default choices made for users (shown in Figure 12). First, having to do with the security of logins, by default, Twitter does not verify login requests using a two-step authentication process (such as requiring a user input a pin number sent to the user’s cellphone before a user can be fully logged in). Further, Twitter does not require personal information to reset a user’s

password by default. Under the header of "Privacy," there are several default settings that impact how user generated information flows to others. First, Twitter (similar to Facebook) allows users to be tagged in photos that have been uploaded to the site. By default, users can be tagged in photos by anyone. Next, by default, user Tweets are not "protected." This means that when a user generates Tweets, those Tweets will be made publicly available, unless the user changes the default position so that Tweets are only accessible that user's followers. Liu, Kliman-Silver, and Mislove (2014) estimated that, in 2007, the percentage of users with protected accounts was as high as 15%, with that number having dropped to less than 5% by 2014. Further, by default, geolocation information (such as exact longitude/latitude) is not made available on Tweets. A user must opt-in to including a specific location within the context of a Tweet. Next, by default, users can be looked up in the Twitter service through the e-mail they registered with. Next, similar to the registration screen, there is a control for "Tailor Twitter based on my recent website visits" that reflects the position selected during registration. Finally, by default, there is a setting called "Tailor ads based on information shared by ad partners" which is on by default. In describing this setting, Twitter states:

> We work with ads partners to bring you more useful and interesting advertising content. We may do this based on information that our ads partners share with us. We hope that this increases the usefulness of Twitter Ads for you. Here's one way it would work. Let's say a flower shop wants to advertise a Valentine's Day special on Twitter. They'd prefer to show their ad to floral enthusiasts who subscribe to their newsletter. To get the special offer to those people, who are also on Twitter, the shop may share with us an unreadable scramble (called a hash) of emails from their mailing list. We can then match that to a hash of emails that our users have associated with their accounts in order to show them a Promoted Tweet for the Valentine's Day deal on Twitter. Another way this works is when a person visits the flower shop's website. In that case, the shop may share with us browser-related information (a browser cookie ID) that we can then match to an account that may receive the Valentine's Day offer. (twitter.com, 2014ac)

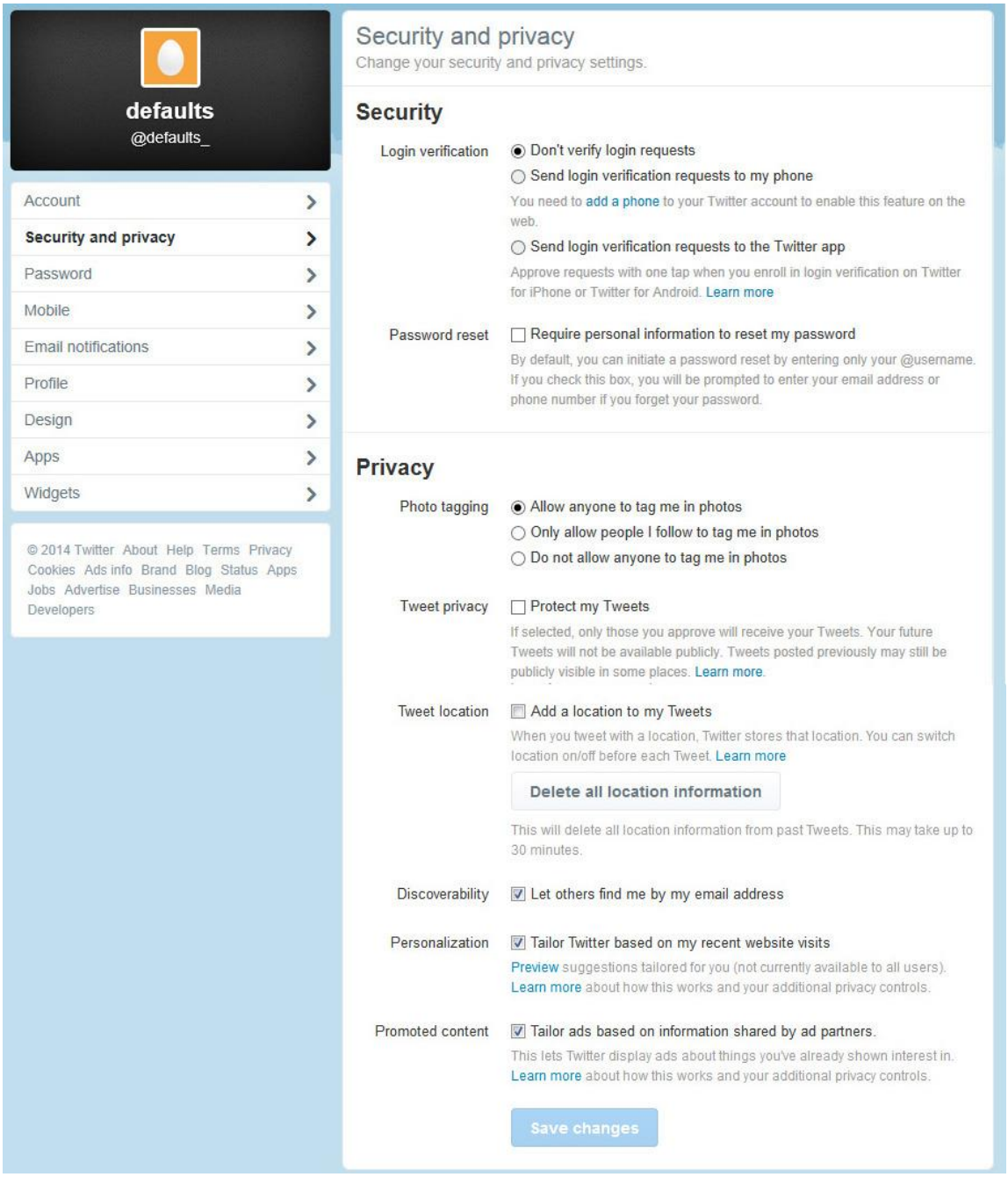


Figure 12. The user "Security and Privacy" settings page.

The "Password" settings page only allows a user to change his or her password or to recover their current one. There are no settings that involve defaults on this page.

The "Email notifications" settings page allows a user to control the kinds of emails that they receive from Twitter. By default, all of the "events" that could trigger an

e-mail from Twitter are selected as active. The full list of all of these events can be seen in Figure 13.

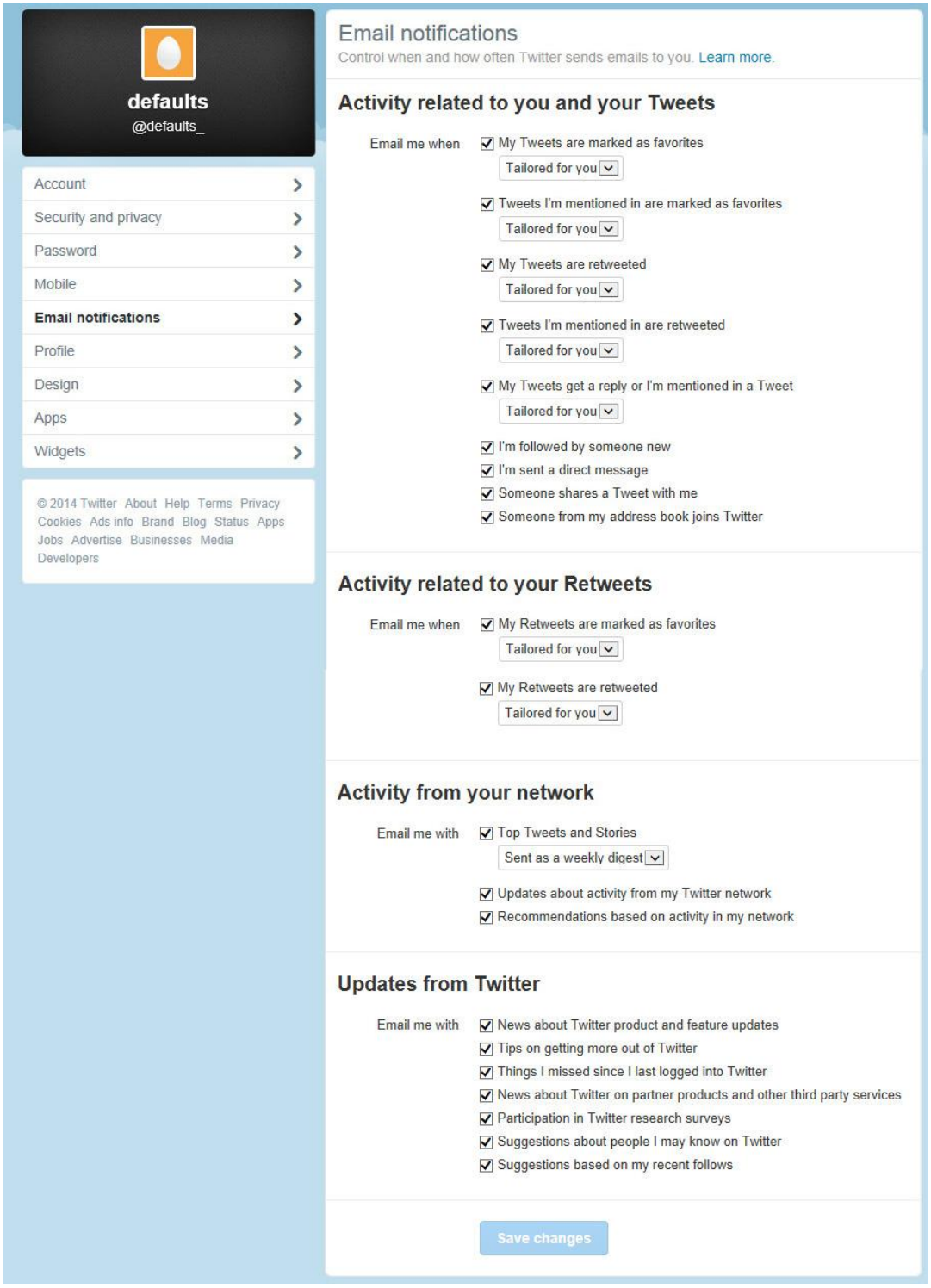


Figure 13. The user "Email notifications" settings page.

On the “Profile” settings page, a user can change the information about themselves that appears on their profile. Other than the username given when a user registers, profile information such as “Location,” “Website,” “Bio” and “Facebook” information is left blank until a user sets it. This is shown in Figure 14.

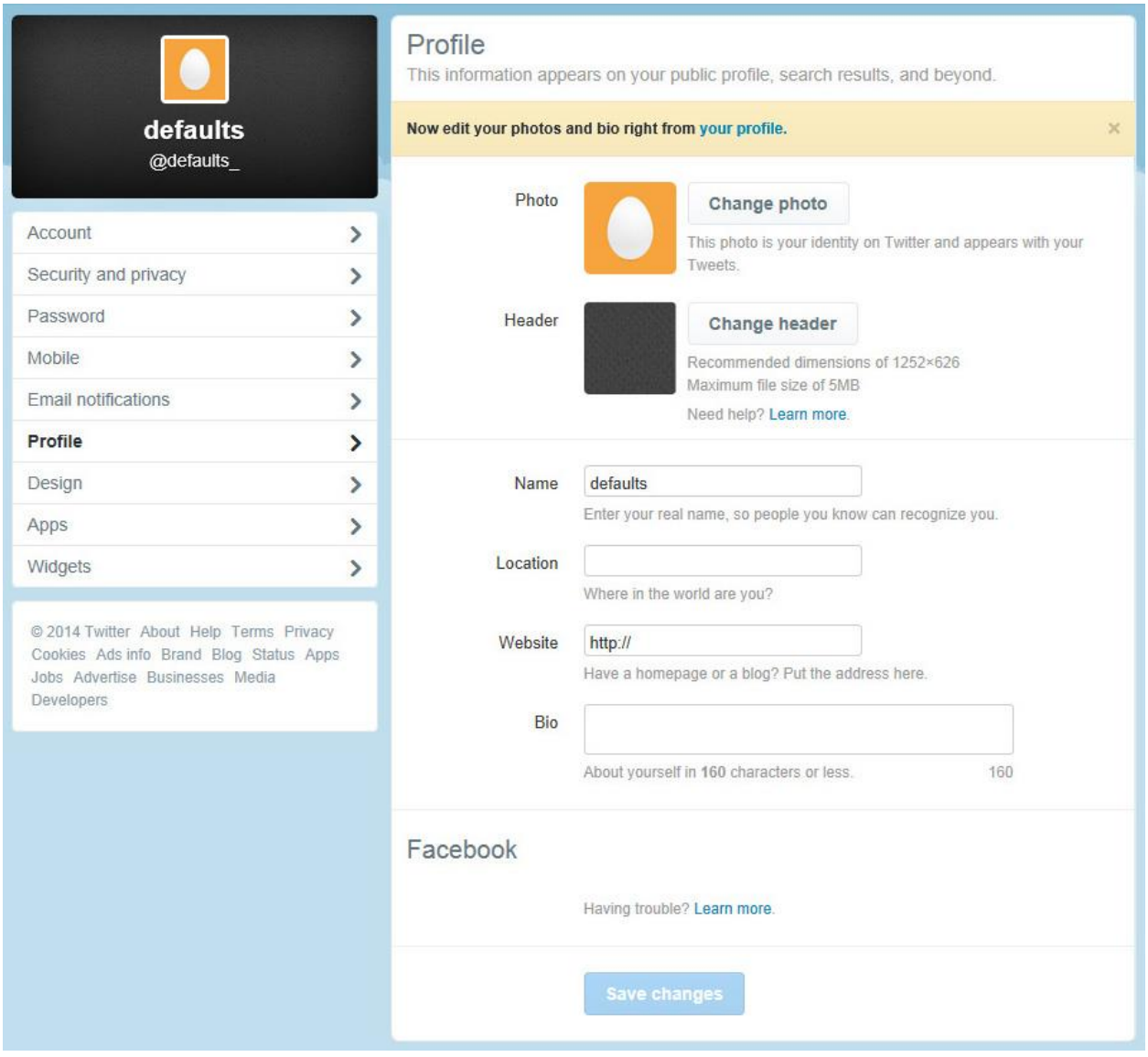


Figure 14. The user “Profile” settings page.

Similarly, the “Design” settings page section simply provides a default aesthetic scheme for the user interface unless changed. This is shown in Figure 15.

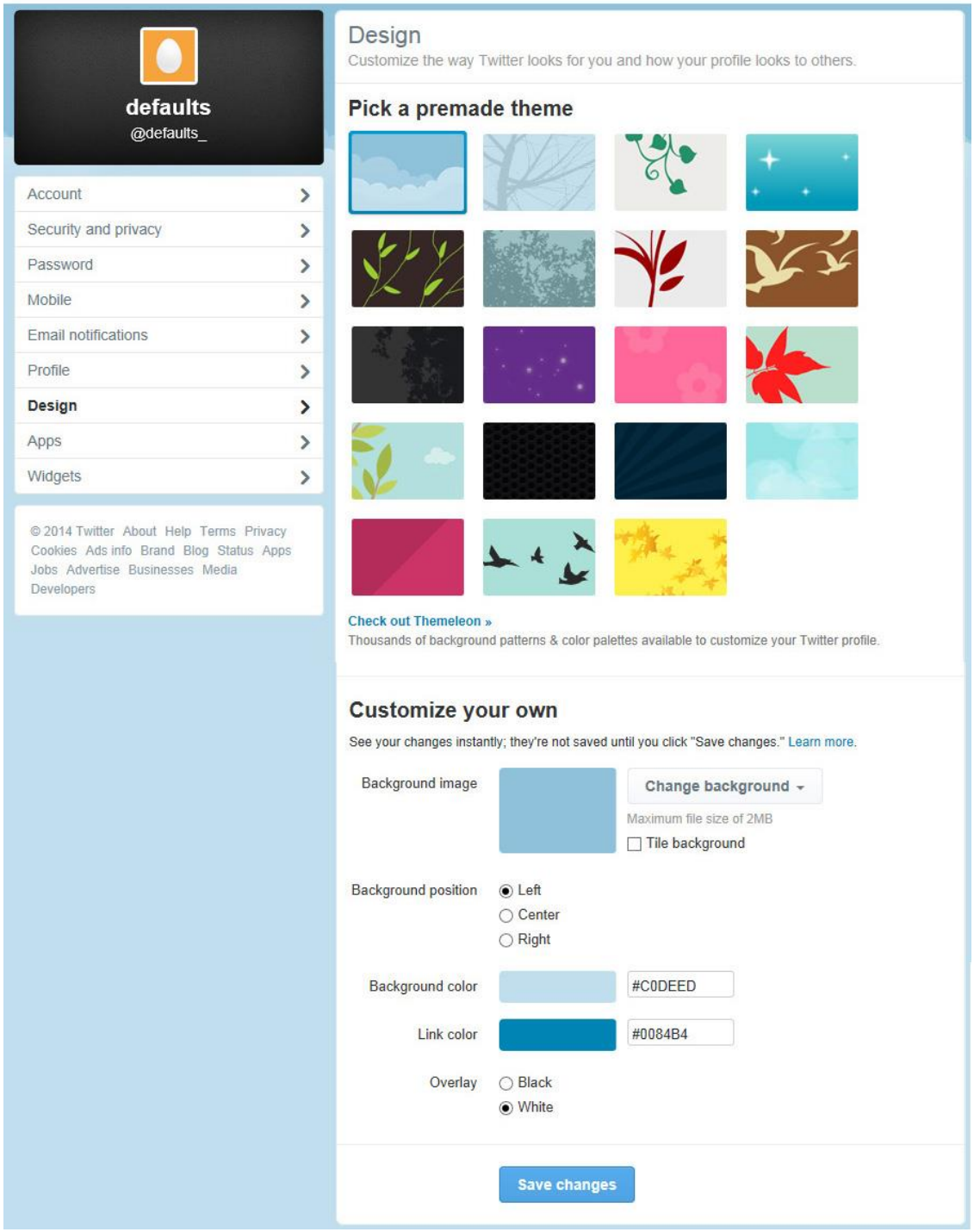


Figure 15. The user "Design" settings page.

Finally, the "Apps" and "Widgets" pages contain no default settings until a user approves applications or widgets as part of their use of Twitter.

Protocols.

There are numerous protocols on the Twitter platform that act as a set of "scripts" for user behaviors within each interface. These scripts govern the range of actions users can perform within the Twitter ecosystem. On the APIs, for example, users are limited in the types of data they can request and post to Twitter by the methods of each API (see Appendix 1 for a list of these methods). Because protocols govern the kinds of actions users can engage in their information production and consumption, they strongly impact and shape information flows within Twitter. This section will focus closely on the protocols that are perhaps the most well-known from the Twitter web-interface: "Tweeting," "Retweeting," "Replying," "Following," "Favoriting," and "Direct Messaging."

Tweeting.

Tweeting is the user-action that is most perhaps commonly associated with Twitter. As discussed briefly in the section on the interfaces, users can create text-messages of up to 140 characters in length and post them to Twitter. As described by Java, Song, Finin and Tseng (2007), tweeting "provide[s] a light-weight, easy form of communication that enables users to broadcast and share information about their activities, opinions and status" (p. 56). Within the scope of the 140 characters, beyond "their activities, opinions and status" users can include URL links to other websites, mentions of other users (signified by typing the "@" symbol and then the other users' Twitter handle) and/or hashtags, which as Small, Kasianovitz, Blanford and Celaya (2012) observe, "function as a folksonomic keyword system for organizing topic-based posts" (p. 176). Although historically the content of Tweets has only been text, more

recently, Twitter has added the ability to share media (such as photos) and geo-location information as part of a Tweet. As of October 3rd, 2013, Twitter indicated in its Securities and Exchange Commission filings that, "Since the first Tweet, our users have created over 300 billion Tweets" (Twitter, Inc., 2013). Interestingly however, it is estimated that up to 40% of registered Twitter users do not tweet (Statistic Brain, 2014).

Retweeting.

Registered users also have the ability to "retweet" the Tweets created by other users. Twitter describes retweeting as the practice of, "re-posting of someone else's Tweet" noting that "Twitter's Retweet feature helps you and others quickly share that Tweet with all of your followers" (twitter.com, 2014r). To illustrate, User A sends a Tweet, and perhaps User B finds that message interesting. User B can choose to "retweet" that message, thereby rebroadcasting the message sent by User A within user B's timeline. Originally—like the use of hashtags and the "@" system for replies and mentions—retweeting was an informal convention developed by users (boyd, Golder, & Lotan, 2010). In the informal model of retweeting, a user would copy the text from another user's Tweet and manually paste it into a new Tweet, adding the prefix of "RT @[UsernameBeingRetweeted]" to the text. This informal user practice was formally codified into a protocol in 2009 when Twitter added a "Retweet" button that now appears under each Tweet in the web-interface (Stone, 2009d). Figure 16 illustrates how a retweeted Tweet appears in the web-interface after Twitter's retweet protocol has been used.

Figure 16. A retweeted Tweet.

Despite the codification of the retweet practice into a formal protocol, some users still rely on the text-based convention for retweeting.

The retweet protocol is a significant part of how information propagates to different users of Twitter. In a study of retweeting practices on Twitter, Suh, Hong, Pirolli and Chi (2010) found that, in their sample of 74 million Tweets from Twitter, 11% of all Tweets were retweets. This number appears to be on an upward trend, as a later study by Liu, Kliman-Silver, and Mislove (2014) has estimated that over 26% of all Tweets are retweets. Other studies have observed that users have many different motivations for retweeting content, including:

> …to amplify or spread Tweets to new audiences… to entertain or inform a specific audience, or as an act of curation… to comment on someone's tweet by retweeting and adding new content, often to begin a conversation… to make one's presence as a listener visible… to publicly agree with someone… to validate others' thoughts… as an act of friendship, loyalty, or homage by drawing attention… sometimes via a retweet request… to recognize or refer to less popular people or less visible content… to gain, either to gain followers or reciprocity from more visible participants… or to save Tweets for future personal access. (boyd et al., 2010, pp. 6–7).

Replies.

Replies are a way for users to respond to another user through a Tweet, thereby creating a linked conversation. In describing replies, Twitter states "An @reply is any update posted by clicking the Reply button on a Tweet. Any Tweet that is an @reply to you begins with your username and will show up in your Mentions tab on the

Notifications page" (twitter.com, 2014ab). Similar to retweets and hashtags, replies began as a user-generated convention. In explaining why Twitter introduced the "reply" button as a formal protocol, Twitter founder Biz Stone stated in the Twitter developer blog:

> The @Replies feature was introduced because we noticed lots of folks putting the @ symbol in front of Twitter usernames as a way of addressing one another. For example: @biz what are you drinking in your avatar? (It's a soy latte.) So, we started linking the @username references and collecting any Tweets that began with @username on one page to make them easier to track. (Stone, 2009c)

Twitter co-founder Evan Williams (2008) has stated that, "Today, @replies are a critical part of how Twitter works" (para. 5). A 2014 study estimated that just under 25% of all Tweets on Twitter are replies (Liu et al., 2014).

Favorites.

Twitter describes "Favorites" by stating, "Favorites, represented by a small star icon next to a Tweet, are most commonly used when users like a Tweet. Favoriting a Tweet can let the original poster know that you liked their Tweet, or you can save the Tweet for later" (twitter.com, 2014u). Twitter maintains a running count of the number of times that each Tweet has been marked by other users as a "favorite," and a list of a user's "favorite" Tweets are publicly accessible (unless that user has marked themselves as having a "protected account"). A study by Suh, Hong, Pirollo and Chi (2010) found that most registered Twitter users do not use the favorites feature. In their study of over 74 million Tweets, the authors found that, "42.5% of Tweets are coming from users with no favorited items... 92.8% of Tweets are coming from Twitter users with less than 100 favorite items" (p. 7).

Following and followers.

The "following" protocol on Twitter is a major driver for the way that information flows between users on Twitter. In describing "following" and "followers" Twitter indicates:

> Following someone on Twitter means: You are subscribing to their Tweets as a follower. Their updates will appear in your Home tab. That person is able to send you direct messages. Followers are people who receive your Tweets. If someone follows you: They'll show up in your followers list. They'll see your Tweets in their home timeline whenever they log in to Twitter. You can send them direct messages. (twitter.com, 2014q)

While theoretically anyone can access another user's public Tweets by visiting that user's page, by following a user, that user's Tweets and retweets are automatically inserted into the followee's timeline. Kwak, Lee, Park, and Moon (2010) describing the practice this way:

> Twitter users follow others or are followed. Unlike on most online social networking sites, such as Facebook or MySpace, the relationship of following and being followed requires no reciprocation. A user can follow any other user, and the user being followed need not follow back. Being a follower on Twitter means that the user receives all the messages … from those the user follows. (2010, p. 591)

Direct messages.

The last protocol that this section will mention is the "Direct Message" protocol. Twitter defines direct messaging by stating, "A direct message (DM) is a private message sent via Twitter to one of your followers. You can only send a direct message to a user who is following you; you can only receive direct messages from users you follow" (twitter.com, 2014w). While replies from users with non-protected accounts can be viewed by anyone, direct messages are usually only accessed and viewed by the author, by the intended recipient, and by Twitter itself.

Algorithms.

Many of the algorithms on Twitter are a part of the site's appeal. For example, Twitter uses an algorithm to parse out the use of hashtags and terms from within Tweets. Twitter then uses another algorithm to determine the most popular hashtags and topics of conversation from this first parsing, which it calls "Trending Topics." The Trending Topics feature "has become an appealing feature for Twitter users, real-time application developers, and social media researchers, thank[s] to the ability to detect trending topics in the earliest stage" (Zubiaga, Spina, Fresno, & Martínez, 2011, p. 2461). However, one of the difficulties of discussing the algorithms present on Twitter is that there is no comprehensive listing of all the algorithms that actually exist within Twitter. Further, while Twitter itself describes some of the algorithms of the site on its help pages in general terms, the inner-workings of the algorithms are also not viewable by the general public. Instead, users (and many researchers) are left guessing what algorithms exist and how they function. This presents a challenge in describing Twitter as a sociotechnical object. As a result, this section will focus on four very prominent algorithms on Twitter whose outputs are made visible to users through the web-interface. These are: the "Who to Follow" algorithm, the "Mentions and Replies" algorithms, the "Trending Topics" algorithm, and the "Tailored Tweets" algorithm. In describing these algorithms, this section will highlight how these algorithms contribute to the informational flows on Twitter.

The "Who to Follow" algorithm is a piece of code that makes suggestions to registered users regarding who they might want to "follow." Twitter (2014y) writes:

> On the left side of your home page, as well as the Notifications and Me pages, you should see a few recommendations of accounts we think you might find interesting in the Who to follow box. These are based on the types of accounts you're already following and who those accounts follow. (para. 7)

Twitter does not provide the exact details how this algorithm works beyond noting generally that it is based on the kinds of accounts the registered user is already following. The "Who to Follow" algorithm is significant to information flows on the platform in the sense that it provides an ever-present (at least through the web-interface) mechanism by which registered users can expand their following network. In expanding their following networks, users grow the volumes of information they "receive" directly within their timelines. As "Who to Follow" does not randomly suggest other registered users, the following network growth achieved through the "Who to Follow" algorithm can be considered as non-random.

Twitter relies heavily on algorithms that help parse and index the 140-character text of Tweets that users upload for different purposes. Two algorithms, the "Replies and Mentions" algorithm and the "Trending Topics" algorithm, are built on top of this parsing. The Replies and Mentions algorithm automatically notifies a registered user if their username appears within any public Tweet on Twitter or within the Tweets of protected users that they follow. It is the "Replies and Mentions" algorithm that allows users to very easily and clearly address each other with an otherwise massive data ecosystem. A user (User A) mentions another user (User B), when they include "@UserB" in the text of their Tweet. Any Tweet that contains "@UserB" anywhere in the Tweet is considered to be a "mention." However, if a Tweet sent by user A begins with "@UserB" it is additionally considered to be a "reply." In both cases, User B will receive notification within the "Notifications" tab of the web-interface. However, there are a number of important idiosyncrasies in how replies and mentions flow to registered users. As part of its FAQ on Replies and Mentions, Twitter notes:

> People will only see others' @replies in their home timeline if they are following both the sender and recipient of the @reply. People will see any mentions posted by someone they follow (all mentions are treated like regular Tweets). People with protected Tweets can only send @replies to their approved followers. If someone sends you an @reply and you are not following the user, the reply will not appear on your Tweets timeline. Instead, the reply will appear in your Mentions tab. (twitter.com, 2014ab)

As a result of this structuring of the Replies and Mentions algorithm, while replies are still "public," they do not automatically flow to all users the same way as other kinds of Tweets.

The "Trending Topics" algorithm is perhaps one of the most well-known algorithms on Twitter. Through the algorithm, users can discover popular temporally bounded topics of discussion within the Twitter ecosystem. Asur, Huberman, Szabo and Wang (2011) note that, "The trending topics, which are shown on the main website, represent those pieces of content that bubble to the surface on Twitter owing to frequent mentions by the community" (p. 2). Trends are determined by the mention of specific hashtags, specific phrases, or specific keywords that have been parsed from the text (this can be seen in the metadata associated with Tweets as part of the APIs). However, there is more than the sheer popularity of a topic that determines whether or not it is displayed to a user. While the exact details about how the "Trending Topics" algorithm functions are unavailable to the public, Twitter does describe the Trending Topics algorithm by stating:

> Trends are determined by an algorithm and, by default, are tailored for you based on who you follow and your location. This algorithm identifies topics that are immediately popular, rather than topics that have been popular for a while or on a daily basis, to help you discover the hottest emerging topics of discussion on Twitter that matter most to you. (twitter.com, 2014s)

What is trending for one individual user may not be what shows as trending for another user in a different location or to a user with a different set of followees. Further, when a user clicks on a trending topic, they are taken to the Twitter search page, which displays the most popular and most recent public Tweets involving that topic. As a result, the Trending Topics algorithm has the potential to shape what information users access outside of their timeline interfaces.

The Tailored Tweets algorithm is the final algorithm this section will mention. Of the four algorithms discussed so far, it is also the newest algorithm on Twitter. This algorithm suggests Tweets and other users that a user might like. However, unlike the "Who to Follow" algorithm, this algorithm uses the user's web history to make suggestions about content. In describing the Tailored Tweets algorithm, Twitter states:

> Tailored suggestions make building a great timeline — filled with Tweets, links, media, and conversations from the people you're interested in — easier and faster. Twitter can now make smarter and more relevant suggestions about who you might enjoy following…. We determine the people you might enjoy following based on your recent visits to websites in the Twitter ecosystem (sites that have integrated Twitter buttons or widgets). Specifically, our feature works by suggesting people who are frequently followed by other Twitter users that visit the same websites. (twitter.com, 2014p)

In this way, the Tailored Tweets algorithm has the potential to shape information flows on the platform by shaping what information users are exposed to within their timeline interfaces.

Data.

Creating a comprehensive account of data and metadata on Twitter is a deceptively difficult task. This is partially a result of the fact that the documentation of data structures Twitter provides is geared towards different audiences (such as web-users versus application developers) who may be using different components of the service

(such as the Web-based Interface and the API Interface). As a result, some documentation regarding data on Twitter is more detailed than other documentation, and no piece of documentation seems to encapsulate and describe all of the different kinds of data that exist on Twitter. For example, Twitter's Terms of Service does not extensively detail the types of data and metadata that exists on the platform, whereas Twitter's API developer guide provides a much more detailed descriptive account of the structure of data and metadata that can be found. This section will proceed by summarizing how the data that exists on Twitter is described in Twitter's Terms of Service, Privacy Policy, and the Twitter Rules, and then how data within Twitter is described in the technical documentation for API developers called the "Field Guide".

Terms of service, privacy policy, and Twitter rules.

Twitter's "Terms of Service" (TOS) is the first of three documents that governs users' access and use of the Twitter platform, and describes some of the data that Twitter maintains within its systems. The TOS refers to the data that users generate as "Content," and notes that "Content" can include things such as "information, text, graphics, photos or other materials uploaded, downloaded or appearing on the Services" (twitter.com, 2014ad). Other than this simple description, the TOS does not provide an account of the exact types of data that exists within the Twitter platform. However, in its TOS, Twitter states:

> Any information that you provide to Twitter is subject to our Privacy Policy, which governs our collection and use of your information. You understand that through your use of the Services you consent to the collection and use (as set forth in the Privacy Policy) of this information, including the transfer of this information to the United States and/or other countries for storage, processing and use by Twitter. (twitter.com, 2014ad)

In Twitter's "Privacy Policy" there is a significantly more detailed account of the type of data that Twitter collects as part of what it calls "Content."

The "Privacy Policy" is the second of three documents that governs users' access and use of the Twitter platform. It articulates *how* Twitter collects data, *what types* of data it collects, and *how it uses* and *with whom it shares* that data. As this is a fairly lengthy list, a detailed chart has been provided in Appendix 2 that traces how the Privacy Policy describes the particular types of data collected, when it is collected, who the data is collected from, what Twitter says the data is used for, what influence users may have on this process (as described in the Privacy Policy), the public/private status of the data (if indicated), whether or not the data is noted as shared with particular users, and any notes about how the data is retained by Twitter. This document reveals a significant amount about not just the different kinds of information that exist on Twitter, but also about how the algorithms, interfaces, protocols, and defaults help give rise to the kinds of information that exist within information flows on Twitter. This section will make only a small number of observations about the types of data that Twitter collects (based on what is stated in the Privacy Policy) rather than trying to summarize every piece of data.

In terms of *what types* of data Twitter collects, the Privacy Policy outlines nine major categories: information collected upon registration; profile information; Tweets, following, lists and other public information; location information; links; cookies; log data; widget data; and information from third-parties. What can be observed about this list broadly is that Twitter collects multiple different types of data that go well beyond the 140-character messages that many think of when they think of the platform. From the description given in the Privacy Policy, it is possible to tell that much of the data that

Twitter receives is information that users must actively choose to disclose, such as profile information or Tweet content. However, some of the data (such as widget data, log data, browser information, cookie information, or data about links a user has clicked) may be collected in an automated fashion that users may not be actively aware of. Twitter also collects information about users not just from users directly, but also from business partners and third-parties, such as from Google Analytics. Further, much of the data that Twitter retains comes to it through the Twitter interfaces, including from the web-interface, the APIs, buttons and widgets, but also from things such as email notifications, applications, and ads.

In their privacy policy, Twitter distinguishes between information that it considers "public" and information that it considers "private." Information such as name and username from the information collected upon registration, profile information, Tweets, following, lists and location information (if a user opts in to including it) is considered "public." Twitter states in regards to this public status:

> Our Services are primarily designed to help you share information with the world. Most of the information you provide us is information you are asking us to make public. This includes not only the messages you Tweet and the metadata provided with Tweets, such as when you Tweeted, but also the lists you create, the people you follow, the Tweets you mark as favorites or Retweet, and many other bits of information that result from your use of the Services. Our default is almost always to make the information you provide public for as long as you do not delete it from Twitter, but we generally give you settings to make the information more private if you want. (twitter.com, 2013e, para. 10)

Importantly however, Twitter still shares information that it considers "private," although this sharing is not "with the world" so to speak. The Privacy Policies state:

> We engage service providers to perform functions and provide services to us in the United States and abroad. We may share your private personal information with such service providers subject to confidentiality obligations consistent with this Privacy Policy, and on the condition that the third parties use your private

> personal data only on our behalf and pursuant to our instructions. (twitter.com, 2013e, para. 20)

The privacy policy also notes that private information can be shared in the event Twitter must comply with a regulation or legal request, or in the event that Twitter is involved in bankruptcy, merger, acquisition, reorganization, or sale of assets.

"The Twitter Rules" are the third of the three documents that governs users' access and use of the Twitter platform. This document governs user behavior on the site as a kind of "rules of conduct." The document itself does not say much about the ontology of data on Twitter; however, as user conduct on the site often involves the creation of data, the rules are important in that they can shape the exact form of information that exists on the platform. Twitter writes that, "there are some limitations on the type of content that can be published with Twitter" (twitter.com, 2014x). Notably, Twitter bans impersonating other people in a way meant to mislead or confuse, infringing on trademarks, publishing or posting other's private and confidential information, posting direct threats of violence, copyright infringement, unlawful uses based on local laws, spamming, phishing, or including pornographic or obscene images in a profile photo, header photo, or user background. As a result, informational content that falls into one or more of these categories may be removed from circulation within the site by Twitter.

While the ontology of data that can mapped from the Terms of Service, Privacy Policy, and the rules governing the content of Tweets may seem as though they are lengthy, they do not actually fully encapsulate all of the different forms of data and metadata that can be found on the Twitter platform. To build a more robust picture of the various forms of data on Twitter, this section next turns to Twitter's technical documentation for its API developers.

Technical documentation of APIs.

While the APIs themselves are discussed in the section on interfaces, what is important for this mapping of the data on Twitter is the documentation of data provided by Twitter in what is called the API "field guide." The Twitter API "field guide" introduces a reader to the various types of data accessible through the APIs in a way akin to the *National Audubon Society Field Guides to Birds.*[23] Twitter states: "Like any ecosystem, the Twitter platform has a variety of flora and fauna. Use this field guide to better understand the most frequently observed wild objects" (twitter.com, 2014a). Twitter describes four classes of data objects API users are likely to encounter: Tweets, Users, Entities, and Places. Each of these types of data is described in greater detail next.

Tweets are the first of four classes of information prominent on the APIs. In describing them as forms of data, Twitter states, "Tweets are the basic atomic building block of all things Twitter. Users tweet Tweets, also known more generically as 'status updates'" (twitter.com, 2014m). There is more to Tweets, however, than just 140 characters. In examining what constitutes a Tweet according to the field guide, one can see that the 140 characters of data is just 1 out of 31 of the fields that can make-up a Tweet; a "Tweet" is actually composed of both data and a significant volume of metadata. Table 2 details all of the different data and metadata fields that can constitute a Tweet.

[23] Complete with illustrations of birds.

Table 2

A “Tweet” and its Associated Metadata.

Field	Description
annotations	Field is currently unused (as of 1/14), noted as "future/beta for status annotations"
contributors	Field indicates users who contributed to the authorship of the tweet, on behalf of the official tweet author.
coordinates	Field can represents the geographic location of this Tweet as reported by the user or client application.
created_at	Field contains UTC time when this Tweet was created.
current_user_retweet	Field details the Tweet ID of the user’s own retweet (if existent) of this Tweet.
entities	Field details entities which have been parsed out of the text of the Tweet. (Such as hashtags, URLs, user-mentions)
favorite_count	Field indicates approximately how many times this Tweet has been "favorited" by Twitter users.
favorited	Field indicates whether this Tweet has been favorited by the authenticating user.
filter_level	Field indicates the maximum value of the filter_level parameter which may be used and still stream this Tweet. So a value of medium will be streamed on none, low, and medium streams.
geo	Deprecated: The"coordinates" field is now used instead.
id	Field contains the integer representation of the unique identifier for this Tweet.
id_str	Field contains the string representation of the unique identifier for this Tweet.
in_reply_to_screen_name	If the represented Tweet is a reply, this field will contain the screen name of the original Tweet’s author.
in_reply_to_status_id	If the represented Tweet is a reply, this field will contain the integer representation of the original Tweet’s ID.
in_reply_to_status_id_str	If the represented Tweet is a reply, this field will contain the string representation of the original Tweet’s ID.
in_reply_to_user_id	If the represented Tweet is a reply, this field will contain the integer representation of the original Tweet’s author ID. This will not necessarily always be the user directly mentioned in the Tweet.
in_reply_to_user_id_str	If the represented Tweet is a reply, this field will contain the string representation of the original Tweet’s author ID. This will not necessarily always be the user directly mentioned in the Tweet.
lang	When present, this field indicates language identifier corresponding to the machine-detected language of the Tweet text, or "und" if no language could be detected.
place	When present, indicates that the tweet is associated (but not necessarily originating from) a Place.

Field	Description
possibly_sensitive	This field only surfaces when a tweet contains a link. The meaning of the field doesn't pertain to the tweet content itself, but instead it is an indicator that the URL contained in the tweet may contain content or media identified as sensitive content.
scopes	A set of key-value pairs indicating the intended contextual delivery of the containing Tweet. Currently used by Twitter's Promoted Products.
retweet_count	Field indicates the number of times this Tweet has been retweeted.
retweeted	Field indicates whether this Tweet has been retweeted by the authenticating user.
retweeted_status	Retweets can be distinguished from typical Tweets by the existence of a retweeted_status attribute. This attribute contains a representation of the *original* Tweet that was retweeted. Note that retweets of retweets do not show representations of the intermediary retweet, but only the original tweet. (Users can also unretweet a retweet they created by deleting their retweet.)
source	Utility used to post the Tweet, as an HTML-formatted string.
text	Field contains the actual 140 character UTF-8 text of the status update.
truncated	Field indicates whether the value of the text parameter was truncated, for example, as a result of a retweet exceeding the 140 character Tweet length.
user	Field contains the user who posted this Tweet.
withheld_copyright	When present and set to "true," it indicates that this piece of content has been withheld due to a DMCA complaint
withheld_in_countries	When present, indicates a list of uppercase two-letter country codes this content is withheld from.
withheld_scope	When present, indicates whether the content being withheld is the "status" or a "user."

While the Privacy Policies describe a set of information associated with Tweets, they do not provide nearly the level of detail about the metadata that surrounds Tweets that the field guide does. Users who actually produce Tweets may be conscious of their 140 character selections, however it remains to be seen as to whether they also are aware of the multitude of metadata that surrounds those messages, as much of this information is produced by algorithms on Twitter's end or is influenced by defaults.

"Users" constitute the second class of data within Twitter that can appear in information flows dealing with the APIs. The Twitter's Developer Field Guide notes, "Users can be anyone or anything. They tweet, follow, create lists, have a home_timeline, can be mentioned, and can be looked up in bulk… Users can be found tweeting, following, and favoriting on Twitter" (twitter.com, 2014ae). Similar to "Tweets," user data objects includes a mix of data and metadata, contain a significant volume of data generated automatically by algorithms, and contain data influenced by defaults. Table 3 details all of the different informational fields that can be found in "user" data-objects.

Table 3

A "User" and its Associated Metadata.

Field	Description
contributors_enabled	Field indicates that the user has an account with "contributor mode" enabled, allowing for Tweets issued by the user to be co-authored by another account. Rarely true.
created_at	Field indicates the UTC datetime that the user account was created on Twitter.
default_profile	When true, indicates that the user has not altered the theme or background of their user profile.
default_profile_image	When true, indicates that the user has not uploaded their own avatar and a default egg avatar is used instead.
description	Field that contains the user-defined string describing their account.
entities	Field contains entities which have been parsed out of the url or description fields defined by the user. (such as URLs, Hashtags, etc.)
favourites_count	Field contains the number of Tweets this user has favorited in the account's lifetime.
follow_request_sent	When true, indicates that the authenticating user has issued a follow request to this protected user account.
following	When true, indicates that the authenticating user is following this user. Some false negatives are possible when set to "false," but these false negatives are increasingly being represented as "null" instead.
followers_count	Field indicates the number of followers this account currently has. Under certain conditions of duress, this field will temporarily indicate "0."
friends_count	Field indicates the number of users this account is following (AKA their "followings"). Under certain conditions of duress, this field will temporarily indicate "0."

Field	Description
geo_enabled	When true, indicates that the user has enabled the possibility of geotagging their Tweets. This field must be true for the current user to attach geographic data when using POST statuses/update.
id	Field contains the integer representation of the unique identifier for this User.
id_str	Field contains the string representation of the unique identifier for this User.
is_translator	When true, field indicates that the user is a participant in Twitter's translator community.
lang	Field contains the BCP 47 code for the user's self-declared user interface language. May or may not have anything to do with the content of their Tweets.
listed_count	Field contains the number of public lists that this user is a member of.
location	Field contains the user-defined location for this account's profile. Not necessarily a location nor parseable. This field will occasionally be fuzzily interpreted by the Search service.
name	Field contains the name of the user, as they've defined it. Not necessarily a person's name. Typically capped at 20 characters, but subject to change.
notifications	Nullable. Deprecated. May incorrectly report "false" at times. Indicates whether the authenticated user has chosen to receive this user's Tweets by SMS.
profile_background_color	The hexadecimal color chosen by the user for their background.
profile_background	A HTTP-based URL pointing to the background image the user has uploaded for their profile.
profile_background_	A HTTPS-based URL pointing to the background image the user has uploaded for their profile.
profile_background_tile	When true, indicates that the user's profile_background_image_url should be tiled when displayed.
profile_banner_url	The HTTPS-based URL pointing to the standard web representation of the user's uploaded profile banner.
profile_image_url	A HTTP-based URL pointing to the user's avatar image.
profile_image_url_https	A HTTPS-based URL pointing to the user's avatar image.
profile_link_color	The hexadecimal color the user has chosen to display links with in their Twitter UI.
profile_sidebar_border_color	The hexadecimal color the user has chosen to display sidebar borders with in their Twitter UI.
profile_sidebar_fill_color	The hexadecimal color the user has chosen to display sidebar backgrounds with in their Twitter UI.
profile_text_color	The hexadecimal color the user has chosen to display text with in their Twitter UI.
profile_use_background_image	When true, indicates the user wants their uploaded background image to be used.

Field	Description
protected	When true, indicates that this user has chosen to protect their Tweets.
screen_name	The screen name, handle, or alias that this user identifies themselves with.
show_all_inline_media	Indicates that the user would like to see media inline. Somewhat disused.
status	Nullable. If possible, the user's most recent tweet or retweet. In some circumstances, this data cannot be provided and this field will be omitted, null, or empty.
statuses_count	Field indicates the number of Tweets (including retweets) issued by the user.
time_zone	Field contains a string describing the Time Zone this user declares themselves within.
url	Field contains A URL (if) provided by the user in association with their profile.
utc_offset	Field contains the offset from GMT/UTC in seconds.
verified	When true, indicates that the user has a verified account.
withheld_in_countries	When present, indicates a textual representation of the two-letter country codes this user is withheld from. See New Withheld Content Fields in API Responses.
withheld_scope	When present, indicates whether the content being withheld is the "status" or a "user."

Entities are the third class of information present within Twitter's APIs. Of entities, Twitter states, "Entities provide metadata and additional contextual information about content posted on Twitter. Entities are never divorced from the content they describe" (twitter.com, 2014e). Practically speaking, entities are pieces of information generated through the algorithmic parsing and processing of Tweets. For example, Entities contain links or hashtags that have been parsed from Tweets. While all of the information within an Entity can be found within its correlated Tweet, entities make it much easier and faster to execute certain algorithms (such as for figuring out which hashtags or media might be trending). The different kinds of entities are detailed in Table 4.

Table 4

An "Entity" and its Associated Metadata.

Field	Description
hashtags	Represents hashtags which have been parsed out of the Tweet text.
media	Represents media elements uploaded with the Tweet.
urls	Represents URLs included in the text of a Tweet or within textual fields of a user object.
user_mentions	Represents other Twitter users mentioned in the text of the Tweet.

Finally, places are the last class of information that Twitter describes in its Field Guide. Twitter states:

> Places are specific, named locations with corresponding geo coordinates. They can be attached to Tweets by specifying a place_id when tweeting. Tweets associated with places are not necessarily issued from that location but could also potentially be about that location. Places can be searched for. Tweets can also be found by place_id. (twitter.com, 2014h)

Places are important to the context of Tweets with geolocation information, as they help identify locations where specific messages are being generated. The data and metadata fields associated with Places are shown in Table 5.

Table 5

A "Place" and its Associated Metadata.

Field	Description
attributes	Contains a hash of variant information about the place. See About Geo Place Attributes.
bounding_box	A bounding box of coordinates which encloses this place.
country	Name of the country containing this place.
country_code	Shortened country code representing the country containing this place.
full_name	Full human-readable representation of the place's name.
id	ID representing this place. Note that this is represented as a string, not an integer.
name	Short human-readable representation of the place's name.
place_type	The type of location represented by this place.
url	URL representing the location of additional place metadata for this place.

Through the terms of service, privacy policy, the Twitter Rules, and through the field guides, this section has provided a descriptive account of the varied types of data that exist within information flows on Twitter. In many cases, this data must be consciously generated by users (such as users actively deciding what characters to put into the 140 characters of a Tweet). However, this review also highlights that much of the data on Twitter (particularly metadata) is generated automatically, influenced by defaults, shaped and produced by algorithms, and may not be readily visible to users of the web-interface. Further, this review shows that some of the data that Twitter maintains (though perhaps not widely accessible) comes from third-parties other than users, such as Google Analytics, or from third-party websites that have Twitter's buttons and widgets installed. While the privacy policy does provide a fairly detailed accounting of the different types of information that Twitter collects, a greater level of granularity and specificity—particularly about the specific forms of metadata that exist on the platform—can be seen in the API Field Guides. Further (and perhaps quite obviously) much of the information that Twitter itself collects and maintains is not "public." For example, Twitter collects log data from users including IP addresses, browser type, operating system, referring web-page, pages visited, location, mobile carrier (if applicable), device IDs, application IDs, search terms used and cookie information, and collects potentially sensitive information about user's travels on the web through its buttons and widgets, but this information is not made available through the web interfaces or the APIs. However, as the privacy policies note, just because user information is not "public" does not mean that Twitter does not share it with some third-parties.

This review of the data on Twitter has been provided not to offer any kind of normative critique of data on the Twitter platform, but to instead point out the large range of data and metadata types on Twitter. Twitter collects data and metadata from a number of sources, and makes data and metadata available to a wide variety of potential receivers, depending on what the data/metadata is and depending on that data's classification of "public" and "private" (a classification done by Twitter). This subsection has attempted to create a map of data on Twitter as it can be established from the governing documents and the API field guide. It is likely that there are data and metadata types on Twitter that this analysis has not adequately captured. However, this highlights the relative complexity of creating a full accounting of the data and metadata structures as part of the overall information flows of the platform.

The techno-cultural dimension: Users.

There are many kinds of users with respect to Twitter. These varied users occupy different parts of the Twitter ecosystem, each acting in different capacities as senders and receivers of information, making use of different parts of the information channels of the technology, and finding interest in different pieces and volumes of informational content. While Twitter touts over 974 million registered users (Koh, 2014), there are potentially even more who make use of the platform. There are users who are nonregistered which simply visit Twitter to read content but not to otherwise contribute; developers and programmers who make use of the APIs to gather data or to build new interfaces; advertisers that "use" the Twitter platform to promote their wares; Twitter's business partners such as Adobe, GNIP, and DataSift who aggregate, process, and/or resell Twitter data; organizations like the Library of Congress who archive Tweets; and even Twitter

itself can be considered as a kind of user. These users engage in quite divergent activities within the Twitter ecosystem, yet all can be considered as "users" of Twitter as they make use of the material technology and the data produced through the technology. This section provides a general overview of some of the different kinds of users of Twitter, making observations about their characteristics as senders and receivers of information, and of their use of the particular parts of the channel of information flows, such as web-interface, the APIs, the buttons & widgets, or other data transmission agreements and data-sharing agreements struck with Twitter. This stratification of users is based on the desire to describe users by the roles they might occupy as senders or receivers of information within the conduit of information flows on the platform, and additionally based on the distinctions between users made by Twitter itself. There are, of course, many alternative ways of conceptualizing users as described in more detail as part of Chapter 2. The differentiation among users provided here is simply a stratification that follows from the desire to talk about different types of information creation and consumption behavior and the use of different parts of the Twitter infrastructure.

Web-interface users.

One of the difficulties in discussing "users" on Twitter is that individuals do not actually have to be registered with Twitter to use many parts of the Twitter interface. For example, anyone who can navigate to the Twitter.com website can see the Tweets that have been generated by registered users with non-protected accounts ("public Tweets").[24]

[24] There is an exception to this: Twitter will block nonregistered and registered users from a given country from gaining access to specific Tweets or the timelines of specific users if Twitter has received a request to block access by an authorized legal entity within said given country. For example, the 5th Criminal Court of Ankara in Turkey (2014) requested that Twitter block access to specific Tweets within Turkey, and as a result, users from Turkey cannot gain access to those messages. Twitter uses a metadata field associated with the Tweets and Users entities to indicate if content should be withheld from any specific country.

These nonregistered users of Twitter can also use the "search" function in Twitter, which allows users to find public Tweets generated by specific individuals or containing specific key terms. Presumably, Twitter collects log data from nonregistered users, although this is not explicitly described in their Privacy Policy. What nonregistered users cannot do, however, is generate Tweets.

Once an individual has "registered" with the site, they are granted the ability to generate Tweets (among many other types of data and metadata) and to act more robustly as a sender and receiver of information. The types of data generated by, and collected from, registered users are detailed in the section on "data." Of the almost 1 billion registered users on Twitter, there is an incredible variety regarding who they are, where they come from, and what their demographic characteristics are. For example, Twitter has been adopted by (and certainly not limited to): students and educators (B. I. Fox & Varadarajan, 2011; Grosseck & Holotescu, 2008; Junco, Elavsky, & Heiberger, 2013; Tiernan, 2013), fans (Highfield, Harrington, & Bruns, 2013; Recuero, Amaral, & Monteiro, 2012), athletes (Hambrick, Simmons, Greenhalgh, & Greenwell, 2010), celebrities (A. Marwick & boyd, 2011), hacking groups such as Anonymous (Mansfield-Devine, 2011), various consumer brands and marketers (Culnan, McHugh, & Zubillaga, 2010; Jansen et al., 2009; Kwon & Sung, 2011; M. Zhang, Jansen, & Chowdhury, 2011), news organizations (Armstrong & Gao, 2010), non-profit groups (Waters & Jamal, 2011), and even politicians and different government agencies (Chi & Yang, 2011; Golbeck et al., 2010; Wigand, 2010). Over 60% of registered Twitter users come from outside of the U.S. (Sanford, 2010). Work done by the Pew Internet and American Life Project has found that 23% of U.S. internet users are Twitter users (Duggan et al., 2015)

and that at least 14% of the overall U.S. adult population are Twitter users (Himelboim, 2014). However, the demographic profile of U.S. Twitter users is not reflective of the full population in the U.S. (Mitchell & Guskin, 2013). Mitchell and Guskin write, "Close to half, 45%, of Twitter news consumers are 18-29 years old. That is more than twice that of the population overall (21%) … Further, just 2% of Twitter news consumers are 65 or older, compared with 18% of the total population" (2013, para. 7). So while there are a large variety of different kinds of individuals who use Twitter in the U.S., this demographic heterogeneity does not match the overall make-up of the U.S. at large.

Among registered users, there are two additional types of users: verified accounts and unverified accounts. Unverified accounts are the "standard" type of account and make up the majority of accounts on Twitter. Typically, Twitter only offers "verified" accounts to celebrities or high profile brands. Accounts that have been verified receive a special blue-icon with a checkmark that appears near their username on their user profile. Twitter describes verified accounts by stating:

> Verification is currently used to establish authenticity of identities of key individuals and brands on Twitter… Twitter verifies accounts on an ongoing basis to make it easier for users to find who they're looking for. We concentrate on highly sought users in music, acting, fashion, government, politics, religion, journalism, media, sports, business, and other key interest areas. We are constantly updating our requirements for verification. Note, verification does not factor in follower count or Tweet count. We do not accept requests for verification from the general public. If you fall under one of the above categories and your Twitter account meets our qualifications for verification, we may reach out to you in the future. (twitter.com, 2014p, paras. 2–3)

As of April 2014, there were 89,000 verified accounts on Twitter, which represents 0.009% of all registered Twitter accounts. Users who are offered verification actually have an additional registration process complete with a tutorial on how to make "good" Tweets that is not offered to unverified registered users (Dash, 2013).

Advertisers.

Advertisers constitute an important category of "users" on the Twitter platform. Advertisers can purchase what are called "promoted Tweets," "promoted trends," or "promoted accounts." The details of each of these products are discussed in greater detail as part of the "Business Models" section of this chapter. Generally speaking, however, advertisers have the opportunity to engage in information flows that have different characteristics than unregistered and registered users. By purchasing a promoted Tweet, promoted trend, or promoted account, an advertiser can have their own content (Tweets, trending topics, or suggestions about who to follow) inserted into the timelines and interfaces of users that may not follow the advertiser's account otherwise. Advertisers essentially have heightened privileges for acting as a sender of information and are able to target specific receivers for this information based on characteristics of those users, such as stated gender, geolocation, or words that those users have used as part of Tweets.

Advertisers on Twitter also receive additional information regarding their advertisements. Through their analytics page, Twitter reports to advertisers that use promoted Tweets the number of times users have been served that Tweet, the number of clicks on a promoted Tweet (and information on which piece of the Tweet users clicked on), the number of times a Tweet has been retweeted, the number of times a promoted Tweet has been replied to, as well as statistics on cost per user engagement. Advertisers who purchase promoted trends are offered information about the users picked up as followers during the campaign period, the number of mentions based on the trend, the number of views of the trend, as well as all of the information offered by the promoted Tweets analytics. Finally, advertisers who purchase promoted accounts are offered

information about the number of times users are served the promoted account, the number of clicks on the account, the number of new followers gained through advertisement, the follow rate during the period and information pertaining to the overall cost of the advertisement (twitter.com, 2014b).

API users.

As of Version 1.1 of the API, users of the API must be registered with Twitter. Non-registered access to the APIs is no longer allowed. However, as noted in the section on the APIs, the kinds of information flows available through the APIs are not uniform among all API users. In order to make use of the "garden-hose," and "firehose" Streaming APIs, API users must have a partnership agreement with Twitter. As a result of these varying levels in API access, the volumetric flow of information from Twitter to the API user is often determined by the receiver's status as either a business partner or non-business partner.

While a full listing of who has access to the "firehose" is beyond the scope of this dissertation, Twitter does make mention of several organizations with which it has data-sharing agreements on its "Twitter Certified" program page. The Twitter Certified program:

> …identifies the best products and services that help businesses thrive on Twitter. The Twitter Certified program consists of two data resellers — licensed to syndicate Tweet content — and a constantly growing list of the leading solutions from the Twitter ecosystem that provide strategic value to Twitter's partners, publishers, and brands. Our Certified product partners deliver valuable functionality beyond what Twitter offers through its own products, through exemplary usage of Twitter's APIs and data products. There are numerous benefits of being a Twitter Certified product partner, including access to Twitter partner engineers, guidance on taking full advantage of Twitter's APIs and data offerings, and exclusive invitations to select beta programs and other Twitter events and activities. (twitter.com, 2014o)

Of the 33 partners listed on the Certified Products page, many of them are noted as having access to the full firehose. For example, a company called BuzzFinder is described as an "analytics tool that draws from the full Japanese Firehose that empowers businesses to understand raw customer feedback, enabling them to recognize new business chances, head off potential incidents, and gain insight into competitors" (twitter.com, 2014f). It should also be noted that two of the certified products, GNIP (which, as of April of 2014, Twitter is in the process of acquiring) and DataSift are authorized to resell historical firehose data, meaning that the flow of information through Twitter's APIs is strictly "real-time," however; historical access can still be acquired from other users in the Twitter ecosystem. It should also be noted here that, although it does not stem from access to the APIs, as noted in the Privacy Policies, Twitter also delivers the full body of all public Tweets to the U.S. Library of Congress (twitter.com, 2013e).

Twitter as user.

It almost appears to be a circular statement to suggest that Twitter should also be conceptualized as a user on Twitter, but Twitter importantly acts as both a receiver and sender of information in relation to the platform. Twitter is a kind of user above all other users, in the sense that it has the widest access to information within the ecosystem, and as it essentially controls the channels of information flow. As noted in the section discussing information flows, it is important to keep in mind that the transmission of information on the platform is a multi-step process. Twitter is *not* being conceptualized here as merely the channel between User A and User B. Instead, when User A creates a Tweet, he or she must first communicate that Tweet to Twitter. Twitter, in this first step,

is acting as receiver. Subsequently, as a second step, Twitter then acts as sender, making that Tweet available to many other potential receivers through the interfaces, APIs and other delivery mechanisms.

The techno-cultural dimension: Informational content.

While the form of informational content available on Twitter has been described at length in the discussion of data and metadata, this section addresses (albeit more abstractly) characteristics of some of the content found on the platform, focusing specifically on Tweets. Registered users have put the technology of Twitter to an incredible variety of uses across various contexts. In their study of genres of Tweets on Twitter, Westman and Freund (2010) identified five common genres of Tweet content: personal updates (for example: "eating a sandwich"), direct dialogue (for example: "@username you should tweet more!"), real-time sharing (for example: breaking news), business broadcasting (for example: "Nike sweatshirts now on sale!") and information seeking (for example: "Can anyone tell me if interstate 794 is still shut-down?"). One of the difficulties presented by this taxonomy, however, is that the categories are not inherently exclusive. For example, the personal update "eating a sandwich" could also be considered real-time sharing as it focuses on present-tense activities. Arguably, it is presence of such "real-time" information that has made Twitter such a phenomenon in the Web 2.0 world. It is what has made Twitter appear as a tap into the *zeitgeist* of the Internet.

For many users, Twitter has become a backchannel or second-screen; a way to participate in real-time conversation while simultaneously watching television (Harrington, Highfield, & Bruns, 2012), while attending conferences (Ebner, Beham,

Costa, & Reinhardt, 2009) or while simply experiencing the mundane activities of everyday life (Miller, 2008). The up-to-the-second nature of the content on Twitter has become a major draw of the platform (Bilton, 2013). In an article in *Time*, Ben Bajarin (2013) explains the importance of Twitter this way:

> For me, it's a real-time news service letting me get instant information, news, events and more in near real time. Throughout the history of broadcast mediums, when major events took place, people would turn to radio or TV to get a sense of all that was happening. I feel that we are in a shift and that Twitter is setting itself up to be the next major broadcast mechanism. (para. 4)

The importance of Twitter for Ben Bajarin among others is inseparable from the fact that the content that Twitter makes available is "real-time." This real-time content has helped spur new uses of the platform. Journalists are now frequently relying on Twitter to identify potential news stories and for source verification (Bennett, 2012). The news organization *CNN* uses algorithmic processing of Tweets to help identify breaking news (Popper, 2014). These receivers of information from Twitter are there because of what can be done with the "real-time" informational content on Twitter. Twitter itself highlights the importance of "real-time" content to the platform in its SEC S-1 Filing when it stated:

> Real-time content allows our users to enhance experiences by digitally connecting to a global conversation as events unfold, and enables our users to engage with each other directly and instantly in the moment and on-the-go. The combination of our tools, technology and format enables our users to quickly create and distribute content globally in real time with 140 keystrokes or the flash of a photo, and the click of a button. The ease with which our users can create content combined with our broad reach results in users often receiving content faster than other forms of media. Additionally, because our platform allows any of our over 215 million MAUs to contribute content, we have a vastly larger production capability than traditional media and news outlets. (Twitter, Inc., 2013, p. 95)

Twitter itself has helped stoke the creation of "real-time" content on the part of users through design elements of the interface and through the implementation of

algorithms that heighten the emphasis on the real-time. For example, the prompt that historically greeted Twitter users, "What are you doing right now?" is a question that invites a response formed to describe the present. Further, in the tutorial offered to newly "verified" users, Twitter suggests that other users are more likely to find Tweets that remark on the present much more interesting than Tweets that describe something that happened in the past (Dash, 2013). The trending topics algorithm further highlights content within Tweets that is being discussed by a large number of people in the present. The orientation of the timeline with the newest material at the top additionally reinforces an orientation towards the present. While there are many different broad genres of content found within Tweets, the "real-time" genre is perhaps the largest and part of what draws users to Twitter.

The socio-economic dimension: Business model.

Fundamentally, Twitter makes money similarly to how many other social media companies make money: by selling advertising and by selling access to user-generated content. Advertising takes a number of different forms on Twitter. The first form is "promoted Tweets," which are Tweets that are interjected into the targeted users' timelines and contain an ad. Promoted Tweets can be targeted to users based on keywords that those users have used in their previous Tweets, depending on those users' stated interests and gender, based on users' geographic location, based on the type of device a user is using to access Twitter, and based on the user's follower/following network (twitter.com, 2014aa). Promoted Tweets are offered on a cost-per-engagement basis, meaning that Twitter charges an advertiser anytime a user clicks on, retweets,

replies, or favorites a promoted Tweet. Writing for Forbes Magazine, Holmes (2013) remarks on the usefulness of such targeting:

> One of the major rubs with traditional ads is inefficiency. Every time a die-hard Prius owner sees an ad for an F-150 pickup it represents a major waste of his time and Ford's money. With Promoted Tweets, this kind of spillage doesn't have to happen. (para. 6)

In addition to promoted Tweets, Twitter also offers "Promoted Trends." An advertiser who buys a promoted trend gets to have a specific hashtag featured at the top of users' "Trends" list all day. Kafka (2013) observes that the purchase of a promoted trend costs roughly $200,000 a day. Lastly, Twitter offers a "promoted accounts" feature that suggests the account of the purchaser as someone to "follow" to other users.

The second mechanism by which Twitter makes money is by selling access to the Streaming APIs, as previously mentioned. Information about the exact cost of access to the full fire hose has been kept well under-wraps, though as noted previously, Google reportedly paid $15 million for access to the full firehose in 2010. However, it is worth noting that in 2011, Twitter did not renew its contract with Google. It is rumored that this was in part because of Google's plans to integrate Tweets into the launch of their social media search products, which may have taken site-traffic and therefore advertising revenue away from Twitter (Constine, 2012).

Both of these mechanisms of revenue generation impact and shape how information flows on the platform. Promoted Tweets and promoted trends create flows of information to users that happen outside of the follower/followee information flow. The selling of access to bulk real-time Tweets further shapes information flows by creating a situation in which only a few are allowed to be receivers of high-volume flows, despite the "public" nature of most Tweets and associated metadata. However, the key element to

the success of Twitter's revenue generation (as with all social media sites) is a reliance on users and user labor to derive economic value. Ads are only valuable if they are seen by users and Twitter can only sell access to Tweets if users are generating Tweets that implicitly have value.

As discussed in Chapter 3, many social media platforms benefit from network effects. In Twitter's SEC filings the business states explicitly:

> Although we do not generate revenue directly from users or platform partners, we benefit from network effects where more activity on Twitter results in the creation and distribution of more content, which attracts more users, platform partners and advertisers, resulting in a virtuous cycle of value creation. (Twitter, Inc., 2013, p. 93).

The number of users on the site relates to the value of Twitter, particularly as those users both consume advertising and generate content. Twitter implicitly highlights the integral role that users play in the part of the SEC filings that identifies the risks that the business faces, stating, "If our users do not continue to contribute content or their contributions are not valuable to other users, we may experience a decline in the number of users accessing our products and services, which could result in the loss of advertisers and revenue" (2013, p. 8). As stated in the section on "informational content" one of the key properties of content on Twitter is that it is often of a real-time nature. Real-time content is one way that contributions become "valuable." As a result of this, users' creation of real-time content is an absolute imperative for Twitter's business model.

John Perry Barlow (1994) wrote, "Most information is like farm produce. Its quality degrades rapidly" (p. 14). Yesterday's news or gossip is not as valuable as today's. Twitter's profitability is steeped in this axiom. However, in order to be a draw as tap into the now and as a source for today's news or gossip, Twitter must recruit as many

users as possible and orient them towards producing this "real-time" informational content. Only by successfully recruiting and orienting this user/labor-base can Twitter become the quintessential tap into *zeitgeist* and medium for participating in up-to-the-minute conversation, and thus be able to generate site-traffic that it can leverage into advertising revenue, in addition to selling access to real-time conversation through its Streaming APIs to third-parties. Perhaps this is why in the orientation for verified users, Twitter emphasizes the role that present tense has over the past tense. However, this begins to point to the important role that the discourse that Twitter generates in shaping user behavior on the platform. By surrounding the platform with messages that help structure individual use towards creating real-time information, Twitter can help to ensure that user labor is oriented towards producing valued and valuable information that then propagates through the information flows of the platform. Functionally then, the success and profitability of Twitter is built in part on its ability to successfully shape user information creation (and hence value extractable from information flows) in a particular way through discourse and the structuring of the site.

The socio-economic dimension: Ownership status.

Van Dijck writes that, "A platform's ownership model is a constitutive element in its functioning as a system of production" (2013, p. 36). Owners of social media have a great deal of control and influence over the way that information flows through the platform. They often establish the "vision" of how a given technology should function and frequently are the ones who make decisions about embracing particular business models, which can further impact how information flows through a given platform. Twitter is no exception to this rule. Twitter's own business history is one marked by

multiple changes in ownership status that have coincided with changes in the way that information flows through the platform.

As recounted by *New York Times* columnist Nick Bilton, Twitter initially began in 2006 as a side-project at the podcasting company Odeo (Bilton, 2013). Bilton attributes the idea for the technology to the product engineers Jack Dorsey and Noah Glass, who then pitched the idea to executives Evan Williams and Biz Stone. The technology was originally conceived of as a site where friends could stream status messages for each other. In 2007, Twitter was spun-off into its own company. Glass was soon pushed out of the new company, Dorsey became Twitter's chief executive, Williams the company's chairman, and Stone became the creative director. The new technology attracted significant venture capital funding, despite the fact that there was not an initial plan for how to capitalize on the technology. According to Bilton, "In exchange for their investment, venture capitalists want[ed], if not a profit, then at least the promise of one eventually" (2013, para. 23). As Twitter took on more and more venture capital, there was increasing pressure on the business' owners to find ways of monetizing the platform. Biz Stone in fact remarked, "Stubborn insistence on a slow and thoughtful approach to monetization—one which puts users first, amplifies existing value, and generates profit has frustrated some Twitter watchers" (Stone, 2010a, para. 1). During this initial growth period, there was also shake-up in management. Dorsey was essentially pushed out of the CEO role, with Williams subsequently taking on the position.

In 2010, there was another shift in management that coincided with changes in the revenue models of the platform. Williams announced that he would be stepping down as CEO to focus on product strategy, while Twitter's former Chief Operating Officer, Dick

Costolo, took over as CEO. Remarking on the change, Williams wrote, "During his year at Twitter, he [Dick Costolo] has been a critical leader in devising and executing our revenue efforts, while simultaneously and effectively making the trains run on time in the office" (Williams, 2010a). Costolo is generally credited for spearheading the implementation of many revenue generation streams on the site today (Isaac, 2013a). For instance, in 2010, Twitter introduced its first advertising system to the platform as a way to generate revenue (Stone, 2010a) and began charging for the use of its data. That year Twitter, "had … revenue of $45 million," (Ante & Das, 2011, para. 4) although it was still operating at a loss due to hosting costs. As noted in the section on business models, the implementation of these two revenue generation models has had a significant impact on information flows within the platform.

In late 2013, Twitter announced that it had filed the necessary paperwork with the U.S. Securities and Exchange Commission to become a publicly traded company (Twitter, Inc., 2013). There was immediate speculation by commenters about what this would mean for the site. Zachary Seward, a writer for the online news outlet *Quartz*, wrote that in response to becoming a publicly traded company, accountable to shareholders, users should expect to see "More ads in more places. More pressure to make more money" (Seward, 2013, para. 21). And indeed, since the IPO there has been an effort to diversify the kinds of advertising offered within Twitter under Costolo's leadership (Shrivastava, 2013). In a 2014 interview, Costolo suggested that, moving forward, Twitter would focus on growing its user-base in order to help revenue. Costolo noted four priorities for gaining and maintaining users: getting new users acquainted with how Twitter works faster, making photos and videos a bigger part of Twitter,

encouraging more conversation among users, and improvements for topic-based discovery of informational content on the platform (Bercovici, 2014). Previously, Costolo has also made a number of remarks regarding the role that discourse plays in gaining users any given platform. In his personal blog, he stated:

> You need to understand what things you are going to do, how you are going to communicate with ALL your customers, etc., in order to maximize the number of new customers that will try your service, and at the same time minimize the number of people who you give a reason to try something else. (Isaac, 2013a)

Interestingly, this in many ways parallels the comments made by Scholz (2008) that were discussed in Chapter 3; that profitability of the Web 2.0 platforms is often dependent on not alienating users with information flows they are uncomfortable with. While Costolo's comments did highlight potential changes to the discourse oriented towards users and some potential changes to the interface, protocols and algorithms, he did not discuss Twitter's profit generation methods. It remains to be seen what the future of commodification of information flows on the site looks like, particularly as it exists in relation to the stockholders and management of Twitter.

The socio-economic dimension: Governance.

There are a number of forces that govern the Twitter ecosystem and can shape or influence information flows. The legal scholar Lawrence Lessig (2006) argues that governance typically occurs through four mechanisms: architecture, the marketplace, social norms, and law. The discussion of data, metadata, algorithms, protocols, interfaces, and defaults earlier in this chapter describes how technical architecture creates a "channel" through which information flows between senders and receivers. This technical architecture governs the ways that various users are capable of interacting with the system. For example, this architecture makes it impossible for users to send Tweets

with more than 140 characters. The discussion of Twitter's business practices and ownership illustrates how Twitter's status as a for-profit business and its position in the marketplace has resulted in a push towards monetizing the platform. This has included actions such as implementing advertisements and selling access to the "firehose" API, changes that have given rise to the current state of information flows on the platform. This chapter has also described, to a lesser degree, how certain social practices on the platform have become implemented into the architecture of the platform. For example, the use of hashtags and stylistic formatting of re-Tweets all began as user-generated conventions. However, so far, this chapter has not addressed the role that law and policy plays in the regulation of information flows on the platform.

Internal policy and external law both exert influence on the informational flows within the platform govern Twitter's ecosystem. For example, Twitter is governed by the laws of the countries in which it operates. These laws can make possible certain kinds of information flows or may restrict certain flows. In the U.S., Section 230 of the Communications Decency Act gives Twitter indemnity from being held legally responsible for what its users say, with the exception of certain kinds of speech associated with criminal activities or the violation of intellectual property rights (Electronic Frontier Foundation, n.d.). As a result, Twitter cannot be held liable for defamation or libel posted by users to the platform. In the absence of this external law (or other laws that govern content), Twitter could have a very different structuring of information flows on Twitter (or perhaps, in the extreme, none at all). However, not all external laws engender information flows. For example, in 2014, a Turkish court asked Twitter to block certain Tweets from being made available to receivers within Turkey, to

which Twitter agreed, while it simultaneously fought the ban in court (Gadde, 2014). Accounting for all of the laws and external forces that are capable of influencing information flows in the Twitter eco-system is beyond the scope of this dissertation. Instead, as van Dijck (2013) does, this dissertation will instead turn to address governance of Twitter by internal forces.

As van Dijck (2013) notes, on social media platforms, governing documents such as terms of service, end-user license agreements, and privacy policies effectively regulate users and information flows in this capacity. These documents serve as legally binding contracts, and as van Dijck (2013) observes, "a site's terms of service are an arena for setting and contending social norms, a struggle that may eventually affect legal rulings" (p. 38). Twitter regulates users through various policy documents, depending on their status as users. For example, registered users are governed by Twitter's "Terms of Service," "Privacy Policies" and "The Twitter Rules;" developers who make use of the APIs are additionally governed by the "Developer Rules of the Road;" and advertisers also have a set of policies that they are expected to abide by (twitter.com, 2014z). These documents set expectations and rules for how users will act in their roles as senders and receivers of information, how they are expected to interact with information channels, and the kinds of informational content that they are prohibited from generating. For example, "The Twitter Rules" spell out boundaries around what is considered unacceptable informational content for Twitter users to generate. Spam, abuse, phishing, and malware are all kinds of informational content there are prohibited from the informational flows of Twitter. API users are given strict rules about the redistribution of content on Twitter. They are told they cannot "sell, rent, lease, sublicense, redistribute, or

syndicate access to the Twitter API or Twitter Content to any third party without prior written approval from Twitter" (twitter.com, 2013a, para. 9). Advertisers are also regulated in the kinds of content they can publish in their ads. For example, Twitter prohibits advertisements that are for adult or sexual products or services, drugs or drug paraphernalia, endangered species products, hate content, tobacco, unauthorized ticket sales, weapons, or other content that is prohibited by trade sanctions (twitter.com, 2014z).

In addition to governing various groups of Twitter users, these documents also play an important epistemic function. For instance, they communicate to users the existence of certain information flows, what information Twitter collects about users, and what Twitter does with such information. Importantly, however, these documents and the information flows on Twitter are not static entities. Instead, the governing documents have been re-written and revised many times in Twitter's history in tandem with changes to the web-interface, the APIs, and the algorithms that operate underneath the surface of the interface. Often times, these governing documents are changed in relation to new types of data being collected, in tandem with the changes in business models (such as when advertising was introduced), and in relation to new information flows on the platform (such as the flow of information to the Library of Congress). Importantly, the "Terms of Service" note that "The Services that Twitter provides are always evolving and the form and nature of the Services that Twitter provides may change from time to time without prior notice to you" (twitter.com, 2014ad, para. 5). While these documents may help users develop an understanding of certain information flows on the platform, and expectations for their behavior and use of the information flows, the documents are ever changing and do not promise a fully detailed account of the information flows of the

platform. While users may be governed by these documents, they cannot count on them to fully disclose the inner-workings of the platform.

Conclusion

To summarize, this chapter began by first defining information flows based on a transmission model of communication that considers 1) the means/mechanism/channel by which information flows, 2) the informational content, 3) the sender, and 4) the potential receivers. Next, the chapter argued that this transmission model can be expanded through the application of van Dijck's (2013) analytical framework for examining social media. Van Dijck's framework considers the constitution of social media sites as techno-cultural elements, such as data/metadata structures, algorithms, protocols, interfaces, defaults, users, informational content; and the socio-economic elements of the platform, such as business model, ownership status, and governance. The chapter argued that van Dijck's framework can be mapped on to the transmission model of communication, such that data/metadata, algorithms, protocols, interfaces, and defaults constitute the "channel;" users constitute senders and receivers, informational content translates as the information; and then the socio-economic elements of business models, ownership status, and governance provide the context that gives rise to the arrangements of techno-cultural elements.

The chapter then proceeded to detail each of these elements with respect to Twitter through a technical reading of the platform, thus providing a baseline account of information flows on Twitter. To use a metaphor, if one conceptualizes information flow as a river, through a close reading of the Twitter APIs and supporting documentation for application programmers, the Twitter for business webpages, Twitter's web interface,

Twitter's SEC filings, the Twitter blog, the Twitter developer forms, Twitter's policy documents, and other secondary sources, this chapter described the makeup of the riverbed. The analysis showed that the facets of information flows of this platform are complex and varied, comprised by numerous individual data and metadata structures, algorithms, protocols, interfaces, and defaults. There are also many different types of users present on the platform, including unregistered users, registered unverified users, registered verified users, advertisers, different stratifications of API programmers and developers, and of course, Twitter itself, who can each function as senders and receivers of information in relation to specific configurations of the information channel. Further, there is a large volume of informational content on Twitter, and this informational content can be categorized into a number of different taxonomies. This chapter also highlighted a number of socio-economic dimensions that influence and shape information flows on the Twitter platform. In particular, these are the for-profit nature of the business models of the platform, the platform's ownership model, and its internal governance.

Moving into the next chapters of this dissertation, this account will serve as a baseline state against which users' *principles-knowledge* of information flows will be surveyed, and against which particular aspects of the technological discourse surrounding Twitter will be compared. The analysis provided in this chapter moves this dissertation closer to answering the primary research question: "In the user-Twitter relationship, what is the state of Twitter users' informational power in regards to the informational flows of the platform? What knowledge of information flows do users have and how does the technological discourse surrounding the site created by Twitter's business purveyors

describe information flows and potentially impact users' knowledge (and hence informational power)?

Chapter 5: Methodology

Introduction

To summarize the argument of the dissertation thus far, Chapter 2 begins by positing a theoretical lens for conceptualizing the relationship between a user and technology, arguing that the relationship inherently involves the negotiation of power. A particularly important facet of negotiated power is what Braman (2006) identifies as informational power; a type of power that serves as an informational base for other forms of power. The second chapter argues informational power is based in part on knowledge of how a given technology functions, what Rogers (2003) refers to as *principles-knowledge* of the technology. Individuals may develop *principles-knowledge* as a technology diffuses throughout society. However, when it occurs, the development of this knowledge can be influenced by factors external to the individual, such as the material design of a technology and the discourse that surrounds the technology.

The third chapter introduces a specific type of technology, social media sites, and surveys how scholars have conceptualized the user-SMS relationship and the operation of power in this relationship. Chapter 3 argues that one of the problems scholars consistently identify in the user-SMS relationship deals with users' lack of power in relation to information flows on SMSs. However, the relevant literature only tacitly recognizes this as a problem involving informational power. Further, the work that empirically explores individuals' knowledge of information flows on social media rarely additionally explores factors that may contribute to this knowledge, such as discourse. Chapter 3 suggests that, in order to address these gaps, further study into the interconnection of individual knowledge, discourse, and informational power should be undertaken and proposes the

popular social media site Twitter as an ideal venue in which to study these interconnections. After reviewing some of the extant research on the user-Twitter relationship, the chapter concludes by presenting the primary research questions of the dissertation: "In the user-Twitter relationship, what is the state of Twitter users' informational power in regards to the informational flows of the platform? What knowledge of information flows do users have and how does the technological discourse surrounding the site created by Twitter's business purveyors describe information flows?" However, to be able to address this question, this dissertation needed to first provide a descriptive account of what the information flows on Twitter are.

Chapter 4 provides such an accounting of the information flows on Twitter through the application of the analytical framework put forth by van Dijck (2013) in her book *The Culture of Connectivity* as part of a technical reading of the platform. The chapter explicates information flows on Twitter by breaking them down into techno-cultural elements, such as data/metadata structures, algorithms, protocols, interfaces, defaults, users, informational content; and socio-economic elements of the platform, such as business model, ownership status, and governance. Through a close reading of the Twitter APIs and supporting documentation for application programmers, the Twitter for business webpages, Twitter's web interface, Twitter's SEC filings, the Twitter blog, the Twitter developer forms, Twitter's policy documents, and other secondary sources, this chapter articulates how the social and the technical comprise information flows on the system, thus providing a descriptive account of the elements of the foundation that make up information flows on the platform. While Chapter 4 analyzes each of the techno-

cultural and the socio-economic facets, it should be kept in mind that information flows on Twitter are only constituted by their operation in tandem.

In this chapter—Chapter 5—the overarching research question is operationalized into three sub-questions:

- Research Question 1 (RQ1): What knowledge of information flows do users have?
- Research Question 2 (RQ2): How does the technological discourse surrounding the site created by Twitter's business purveyors describe information flows?
- Research Question 3 (RQ3): In the user-Twitter relationship, what is the state of Twitter users' informational power in regards to the informational flows of the platform?

This chapter details the methodological approaches used to address these research questions.

Research Question 1 is addressed through the descriptive and inferential statistical analysis of a quantitative self-administered user survey distributed to over 15000 students, faculty, and staff from a large public, urban Midwestern university. This survey asked participants questions about demographic information, such as age, gender, education; about their use history with Twitter; and asked a series of technical questions about information flows on the Twitter platform developed from the descriptive analysis of information flows on Twitter from Chapter 4. In the context of the user-Twitter relationship, these individuals' understandings of the information flows of the platform contribute to what makes up the individuals' informational power relative to the technology. The results of the analysis are reported in Chapter 6.

Research Question 2 is addressed through a critical discourse analysis of the descriptive language Twitter, Inc. presents users during the registration process. This includes the Twitter.com landing page; the "Join Twitter Today" page; the policy documents, including the Terms of Service, Privacy Policy, The Twitter Rules, and the Cookie Use statement; the new-user orientation tutorial; and the Twitter.com web-interface. The analysis focuses on examining the characteristics and themes of how information flows are described in these texts. By inductively analyzing how the information flows on Twitter are described in this discourse, this study explores how these depictions could further or might hinder the development of users' *principles-knowledge,* and hence informational power. The findings from this discourse analysis are presented in Chapter 7.

Chapter 8 synthesizes the findings of the survey and of the discourse analysis, along with the technical analysis of information flows present on Twitter as found in Chapter 4. In doing so, this final chapter addresses the overarching question, RQ3: "In the user-Twitter relationship, what is the state of Twitter users' informational power in regards to the informational flows of the platform?"

Individually, the findings from each part of this study represent important first-steps in filling a gap in the scholarly work to date on Twitter. However, it is together that they contribute to a broader discussion about individual informational power in relation to Twitter. By using a traditional survey approach alongside a critical discourse analysis to probe the state of individual informational power, this project contributes new knowledge to the study of user informational power in the contemporary Web 2.0

landscape while simultaneously creating a unique departure point for future research projects.

Each operationalized research question requires different data to answer and a different plan of study for investigation. The next two sections of this paper outline the methodological approach used to address RQ1 and RQ2, respectively. Each section is broken down into six component parts: 1) an introduction that includes a description of the data needed to satisfy the research question; 2) information about the sources that the data was gathered from, including information on how the data was sampled, the location and method by which the data was gathered, and other general notes about methodological considerations; 3) a description of the instruments or measurements that were used to gather the data, including justifications as to the appropriateness of the tool; 4) a description of the procedures of how the instrument was applied to the data and/or how the data collection instrument was administered; 5) a description of how the data was analyzed, including a justification of why such particular analysis methods are appropriate; and 6) a description of the delimitations (where limits have been set by the researcher) and limitations (which the researcher cannot control) of each approach. The overall goal of this chapter is to describe how this project gathered data, analyzed it, and produced the findings that address RQ1 and RQ2 in Chapters 6 and 7, thereby setting the stage to address the larger question about individual power (RQ3) addressed in Chapter 8.

RQ1: What knowledge of information flows do users have?

Introduction.

To address Research Question 1, this study elicited users' knowledge of the information flows on Twitter through an exploratory survey and then applied descriptive

and inferential statistical tools to explore the responses. As there are no other studies that have surveyed Twitter users' knowledge of information flows on the platform, an exploratory approach was the most appropriate method to answer this question. In exploratory surveys, the research question remains open-ended and there is not a specifically testable hypothesis (Adams, 1989). Instead, from the initial investigation, specific hypotheses may surface that could serve as a direction for future research.

Information sources: population.

U.S. based Twitter users who are above the age of 18 are the population of interest for this study. One of the significant challenges in studying any population is sampling it in a sound and representative manner so that findings can be generalized back to the larger population. Unfortunately, in studying this particular population, true "random" sampling is extremely difficult and not possible within the constraints of this project. There are multiple reasons why this is the case. First, there is no overall list of U.S. Twitter users over the age of 18 that can be randomly sampled from. Second, Twitter does not use a sequential id numbering system for users, so generating random number strings in hopes they would correlate with user IDs is not an option. Third, while it is theoretically possible to take a random sample from the full public stream of Tweets through the APIs, this would bias any sample towards users who have tweeted recently. Further, contacting hundreds of users whose usernames were pulled from the APIs would potentially violate Twitter's Terms of Service.[25] While there are studies that use random samples of Twitter users, such as those done by the Pew Internet and American Life

[25] An approach such as this was tried by Watson (2012), and it resulted in Twitter banning the account used to contact Twitter users 4 times, even after attempts were made to comply with the spirit of Twitter's TOS.

project, the cost associated with the data collection methods they use[26] is somewhere between $40k-$50k and remains (unfortunately) well beyond the means of this project. As this is an exploratory research study, this project instead relies on a purposive random sampling of individuals from the population of students, staff, and faculty at a large public, urban Midwestern university. This does raise issues for the generalizability of the findings, which is addressed in more detail in the *Delimitations and limitations* section. Information about how this population was randomly sampled and contacted is detailed in the "Procedures" section.

Description of the instrument and measurements.

In order to explore users' knowledge of informational flows on Twitter, data were collected using a self-administered Qualtrics web survey. The survey contains a total of 56 questions split over 13 pages. The first page of the survey informs participants about the study, the study's purpose, their rights as study participants, how data will be collected, stored and protected, provides information about the IRB approval of this project, and asks whether they agree or do not agree to take part in the survey. The second page of the survey asks participants for basic demographic information, such as age, gender, education, and whether they have never used Twitter, have used the Twitter website but do not have a registered an account, have a registered account on Twitter, or formerly had an account on Twitter, but deleted it. The third page of the survey—which is only shown to participants who indicate they have used the Twitter website but have not registered an account, have a registered account on Twitter, or formerly had an account on Twitter, but deleted it—asks about the means by which they have accessed Twitter, whether or not they have ever sent a Tweet, how long ago they last posted a

[26] Random digital dialed phone surveys completed by a third-party data collection agency.

Tweet to Twitter, the last time they visited the Twitter.com homepage, whether they characterize their use of Twitter as "almost never use it," "occasionally use it," "use it fairly regularly," or "use it often," and how long ago they first signed up for Twitter. The next nine pages of the survey probes respondents knowledge of particular aspects of information flows on Twitter, including the types of data/metadata collected on Twitter, how protocols shape information flows on Twitter, how algorithms shape information flows on Twitter, about the default settings on Twitter that shape what information is collected and/or displayed, about the different interfaces on Twitter, about Twitter's business models, about the governance of Twitter, about the different types of users on Twitter, and about the ownership of Twitter. As detailed more fully in the section on instrument development, these questions were developed based on the analysis conducted in Chapter 4. With the exception of the page contain the question on the business models of Twitter, participants were presented with either true or false statements about information flows on Twitter and were asked to indicate whether, "Yes, this is correct," "No, this is incorrect," or "I don't know the answer to this question." For example, as part of the "data" page, participants were given the (inaccurate) statement, "Messages on Twitter (also called 'Tweets') are limited to 210-characters in length" (the character limit, as noted in Chapter 4, is actually 140). For the page on Twitter's business practices, participants were asked to identify the ways that Twitter generates revenue from a list of nine possible choices; four true, four false, and an one "I don't know the answer to this question" option. The final page of the survey asked participants whether they feel like Twitter is discussed in the news: "Never," "Occasionally," "Sometimes," or "Regularly;"

how often they keep up with news about Twitter; and about whether they have read Twitter's Terms of Service, Privacy Policies, and "The Twitter Rules."

Participants were allowed to skip any questions they wished to, with the exception of the consent question. To improve the reliability and validity of the instrument, a participant attentiveness question was added to the tenth page of the survey. This reliability question is drawn from a similar one used in the digital-literacy survey developed by Hargittai (2009).

Self-administered surveys that ask participants about their knowledge of technological platforms are a type of instrument often used in digital literacy studies (Hargittai, 2009; Park, 2013) and in the area of privacy studies (Acquisti & Gross, 2006; S. Fox, 2000; Turow, 2003; Turow, Feldman, & Meltzer, 2005). In the development of her digital literacy survey, Hargittai (2009) argues "that the majority of people do not make up their responses to these questions" (p. 130), and thus, the instruments do generally function as a reliable tool for measuring individual knowledge. However, steps to ensure the reliability and validity of the survey were taken to make the instrument as sound as possible. In addition to the attentiveness question, to increase the reliability and validity of this project's survey, the general structure and measurement scales used in this survey instrument were modeled on those used by Hargittai (2009) and Fuchs (2009) to measure user knowledge. As will be discussed in the next section, in order to ensure that problems associated with question wording were not prevalent, this survey went through a multi-step refinement process.

Instrument development.

As the survey instrument used to address RQ1 is an entirely new instrument, to help improve its reliability, it underwent a multi-step refinement process before the survey's initial launch on October 6th, 2014. The digital literacy surveys of Hargittai (2009) were the initial inspiration for the design of this survey. However, after an initial draft of the survey was developed, an informal pilot test revealed a number of shortcomings in the scales used for measurement. To fix these shortcomings, the survey was redesigned based on the layout of the survey used by Fuchs (2009) in his study of student familiarity with organizational surveillance practices on the social media sites studiVZ, MySpace, and Facebook. In this study, Fuchs presented participants with statements about surveillance practices on each of the platforms that are either true or false, and then participants are asked to indicate whether they believe it is a true or false statement. Fuchs then reports the overall percentages of participants who gave correct responses.

The second-draft of the survey instrument used in this dissertation borrowed from the "quiz" like nature of Fuchs' work by presenting a series of true or false statements about particular aspects of information flows on Twitter. After some review, I decided that, in addition to being able to indicate that the statement was either "correct" or "incorrect," participants would also be afforded the opportunity to indicate that they did not know the answer to a question. The reason for this decision was first, a concern that participants would unnecessarily skip questions that they did not know the answer to; second, the realization that a respondent claiming non-knowledge of a particular facet of information flows would perhaps be just as important as a participant indicating a correct or incorrect response; and third, the concern that many respondents would simply guess if

they did not know the answer, thus undermining any potential findings. While it is still possible that respondents may have just guessed about the truth of statements about information flows, they would have less incentive to do so. Further, to break up the flow of the "True/False" nature of the survey, I also decided to provide one question that used a multiple-choice selection, where the selections could be made from a set of true and false statements (the business practices questions).

A test version of the survey was created in the web-survey platform Qualtrics. Online survey consent information provided by the University of Wisconsin-Milwaukee's Institutional Review Board (IRB) was also added to the beginning of the survey to let participants know about the scope of the study, the types of questions that would be asked, how their data would be collected and protected, and their rights as study participants. The test version of this survey was then circulated to six doctoral students within the University of Wisconsin-Milwaukee's School of Information Studies. These students were asked to evaluate the wording of the questions and to test to make sure that there were no problems with the skip-question logic used in the survey. After integrating a half-dozen suggestions on question wording, the survey and plan of study was sent to UWM IRB in late September for human-subject research ethics approval. The UWM IRB approved the study on Sept 30th, 2014. Its approval number is IRB #15.064. A copy of the final survey instrument is included in Appendix C.

Procedures.

The random sample for this population was selected after consultation with a technical support team at the large Midwestern university where the study took place. The technical support team was able to provide an initial randomly selected list of 5000

active email addresses from the public directory of students, faculty, and staff at the university. This was accomplished by retrieving the entire public directory of active e-mail accounts and then randomly selecting 5000 addresses via a Python script. It is possible for individuals at the Midwestern university to opt-out of having their e-mail addresses included in the public directory. As a result, more privacy conscious individuals may not have been included in the sample.

The list of 5,000 e-mail addresses was then loaded into Qualtrics, and at noon on October 6, 2014, the initial sample pool was e-mailed and invited to participate in the web-survey. In addition to the survey link, the e-mail contained information about the project, the goals of the project, indicated what kind of information would be collected and how it would be stored, and provided contact information of the primary investigator and the IRB office that oversaw the approval of the research. A copy of the recruitment e-mail can be found in Appendix D. Individuals were told in the e-mail that they did not need to be Twitter users in order to participate. Individuals who clicked on the survey were taken to a landing page that contained the informed consent information and IRB contact information for the project. After the informed consent process, the individuals who agreed to participate were taken to the survey itself.

As the response rate from this type of "cold-call" outreach effort was expected to be, at best, around 5% - 10% based on the response rate found in a similar recruitment effort used by Vitak (2012), it was determined that a random sample of 5,000 e-mails would be an appropriate amount to yield a result of at least 300 responses. Twenty-four hours after the initial e-mail, Qualtrics indicated that 75 of the 5000 individuals contacted had fully completed the survey. A reminder e-mail was sent to the sample pool after a

one-week period. Twenty-four hours after the first reminder e-mail was sent, a total of 130 persons had completed the survey (a completion rate of less than 3%). After failing to meet the expected response rate, I contacted the technical team at the university and asked for a second additional random sample, pulling an additional 10,000 randomly selected e-mail addresses (after e-mail addresses from the first random pull were excluded). Due to e-mail traffic limits imposed by the university where the survey was distributed, only 10,000 e-mails (including reminders) could be sent to their mail servers a week, and as a result, the second sample of 10,000 individuals was split up into two groups. In the first week the survey was open, 5,000 randomly selected persons were invited to take the survey. In the second week, another 5,000 unique individuals were invited, and in the third week, the last 5,000 unique e-mail addresses were invited. Each group was given a reminder e-mail after one week. The survey was closed after roughly one month. A total of 449 persons fully completed the survey, yielding a completion rate just shy of 3%.

Data analysis.

Once the survey was closed, the responses were exported and analyzed using the SPSS statistical software package. As part of data cleaning, respondents who either failed or did not answer the attentiveness question were removed from the results. A total of 5 respondents failed the attentiveness question and another 10 did not answer the question. This resulted in a total "valid" sample of 434 individuals ($N = 434$). Descriptive statistical analysis was used to summarize the overall responses to each question. Once this initial descriptive work was complete, cross tabs and chi-squared tests were used to explore

differences among respondents in question responses based on demographic characteristics, such as user-status.

Delimitations and limitations.

As with any study that relies on a self-administered web-based survey, there are some limitations and delimitations that must be noted. The first delimitation of this project is found in the nature of the sample. Naturally, it would be a more ideal scenario to be able to acquire a truly random sample of U.S. Twitter users; however, locating such a sample in a cost-effective manner is not within the scope of this project. Instead, this project must rely on a purposive sample of a still quite interesting sub-population. As a result of this sampling, however, there are limitations on the generalizability of the findings. Second, despite work by Hargittai (2009) that suggests knowledge surveys are a reliable tool, self-reporting always carries the potential risk that participants will misrepresent their knowledge. This is a natural limitation of this particular method. To help increase the study's reliability and validity, individuals who failed the attentiveness questions were removed from the results. Despite these limitations, this study can serve as an exploratory launching point for determining whether or not further research—perhaps with greater resources devoted to random-sampling or with specific hypothesis about a population in mind—would be of potential value.

RQ2: How does the technological discourse surrounding the site created by Twitter's business purveyors describe information flows?

Introduction.

As discussed in the review of relevant literature, individuals develop their knowledge about a technology not only through their direct experiences with the

technology, but also through the consumption of messages about the technology. According to Rogers (2003), communications about a given diffusing technology can influence individuals' *awareness-knowledge, how-to knowledge,* and/or *principles-knowledge.* These influential communications can come from a variety of different sources such as friends, family, newspapers, the creators of the technology, etc., and an individual may encounter these communications in a variety of social contexts and situations. However, as the combined works of Pfaffenberger (1992), Rogers (2003), Hull, Lipford and Latulipe (2011), and van Dijck and Nieborg (2009) argue, messages from a technology's creators, designers, and purveyors can play a *particularly* influential role in the formation of individual knowledge regarding a technology. Therefore, to address RQ2 this project explored the language that Twitter's creators, designers, and purveyors presented to individuals signing up for Twitter through the Twitter.com homepage, using the methodological approach of discourse analysis.

For the purposes of this dissertation, with respect to Twitter, the technology's founders, designers, and purveyors were considered as part of one coherent business organization, Twitter, Inc. When Twitter, Inc. creates descriptions of the Twitter platform for public consumption, it presents an account of the technology that is not absent of motive. The language Twitter, Inc. chooses is not simply a transparent window that clearly depicts how the Twitter technology works, what its possibilities for use are, or what is significant about the platform. Instead, this language can be considered as a kind of organizational rhetoric, a type of speech organizations deploy to influence general assumptions held by the public for strategic purposes (Cheney, Christensen, Conrad, & Lair, 2004). Language is a tool that Twitter, Inc. uses (consciously or unconsciously) to

influence the public's general assumptions and knowledge about the platform for a particular set of purposes, for example, convincing people to sign-up. Discourse analysis, as a method, offers a toolkit for unpacking and breaking down this language, making clear that the descriptions that Twitter, Inc. provides are not "simply a neutral means of reflecting or describing the world" (Gill, 2000, p. 172). Instead this use of language can be viewed as purposeful, "performative and functional" (Rapley, 2008, p. 2). In this project, discourse analysis provides an inroad for exploring how this language depicts the Twitter platform in a way that may influence the public's general assumptions about information flows, particularly in ways that may serve Twitter, Inc.'s interests.

Technology companies naturally have motivations for describing their technologies in ways that promote adoption of the technology among the public. This is, after all, how these organizations generate revenue. However, as discussed in Chapter 3, Web 2.0 technologies have additional economic properties beyond those of other kinds of goods. As a result of these economic properties, many Web 2.0 purveyors must dually focus on increasing *both* adoption of the technology among the public, and increasing (or at the very least, maintaining) the levels of use among current users to try to realize profit. As Scholz (2008) observes, user labor is a critical part of the revenue generation processes of most Web 2.0 businesses. As a result, these companies have profit-based motivations for avoiding presenting non-users who may someday contemplate becoming users and current users with language that might give them a reason to choose not to adopt a technology, to slow down in their use of a technology, or to stop their use of a technology. Scholz's work, along with that of Langlois, McKelvey, Elmer and Werbin (2009), suggests that details about how the information users produce flows through a

platform and eventually become commodified may be particularly avoided by Web 2.0 businesses. Hence, the language Web 2.0 businesses generate oriented towards users may not contain descriptions of the platform that fully reveal all aspects of information flows, and as a result, may not promote development of full or robust *principles-knowledge*. Instead, this language may have a more targeted focus on encouraging the development *how-to knowledge*, particularly *how-to knowledge* centered on information production and consumption through the technology.

Twitter, Inc., as a Web 2.0 company, faces this same set of economic pressures. As a result, the descriptions of the platform that Twitter, Inc. produces for users and the wider public may naturally present a particular perspective on informational flows on the platform. The question that this dissertation therefore asks of Twitter, Inc.'s language is: what aspects of the information flows on Twitter does this language draw attention to and what does it gloss over? It follows Twitter, Inc.'s messaging to users can be interrogated to trace how the language promotes specific kinds of knowledge about information flows on the platform. What does this language describe and what does it remain silent about? What kind of image of the information flows on Twitter does this language project? The presence and absence of particular descriptions of information flows in Twitter, Inc.'s language, in addition to the specific characteristics of the information flows that this language depicts, represents an important potential source of influence for the development of an individual's *principles-knowledge,* and hence informational power.

As the analysis in this dissertation explores a part of the discourse generated by Twitter, Inc. in order to examine the way that it promotes certain patterns of belief and knowledge with an eye towards the implications for individual power, this particular

methodological approach can be categorized as a *critical* discourse analysis (CDA) (Fairclough, 1995; Fairclough & Wodak, 1997; van Dijk, 1993, 2003). Critical discourse analysis interprets texts by focusing on how they perpetuate or reproduce ideology, domination, or forms of power imbalance and inequality. In this case, this dissertation explores how Twitter, Inc.'s language use relates to the formation or impedance of individual informational power. Critical discourse analysis, however, does not focus solely on the content of a text (Fairclough, 1995). It also considers the processes of production, distribution, and consumption that form the context in which the texts are embedded. While this chapter has already provided a brief description of the socio-economic context Twitter operates in as a Web 2.0 business, the analysis in Chapter 7 provides additional details about the context that helped shape the form of the texts. Through the exploration of context and content, this research highlights the ways that this discourse may impact users' informational power in relation to the sociotechnical system.

The choices that Twitter, Inc. makes about language use are important not just because of their immediate and direct impacts on individual knowledge, but also because each use and repetition of a particular description makes that text more salient for future uses. As Johnstone (2008) observes:

> [E]ach time a world is created in discourse it becomes easier to create that world again in subsequent discourse. Particular choices can come to stand for whole ways of seeing things, whole ways of being, and those ways of seeing things can come to seem natural, unchallengeable, and right. (p. 46)

The particular ways that Twitter, Inc. describes (and repeatedly describes) the information flows on the Twitter platform impacts the potential for (re)deployment of these descriptions in a variety of new contexts and settings. For example, a newspaper reporter may repeat the language that Twitter, Inc. uses to describe its service when

writing a story about the platform. If the descriptions that Twitter, Inc. creates are overly selective in their descriptions of information flows, there is an even greater potential for this discourse to impact not just individual informational power in the current moment, but in the future as the descriptions are used and re-used. This makes this project's approach to analyzing Twitter, Inc.'s descriptions of information flows all the more necessary.

Information sources.

This study focuses on a specific kind of discourse produced by Twitter, Inc.: messages present on Twitter.com that a user would encounter during the process of registering for an account. This includes the Twitter.com "landing page;" Twitter's "Join Twitter Today" page; Twitter's Terms of Service (ToS), Privacy Policy, the Twitter Rules, and Twitter's Cookie Use statement; Twitter's new user orientation tutorial; and the web interface that new users are brought to once they have completed the sign up process. The text under analysis was captured from the Twitter.com homepage on October 31st, 2013 through a combination of HTML captures and screen-captures.

These texts were selected purposefully, as they are a form of messaging that many in the population of interest (U.S. Twitter users over the age of 18) have likely been exposed to at least once. Despite the fact that these texts engage different genres of writing, they are united by the fact that they all convey messages about the Twitter platform which were generated by Twitter, Inc. In this sense, for the purposes of this dissertation, they are considered a coherent technological discourse generated by Twitter, Inc. that constitutes a narrative about the Twitter platform.

Instruments and measurements.

Bauer (2000) writes: “People use language to present the world as knowledge” (p. 135). It is in this spirit that this project engages discourse analysis to make visible the characteristics of the world of information flows depicted in Twitter, Inc.’s language. Generally speaking, a discourse analysis is a “careful, close reading that moves between text and context to examine the content, organization and functions of discourse” (Gill, 2000, p. 188). Discourse analysis, however, often does not have a set of “hard and fast” rules for how one should go about doing the actual analysis (Rapley, 2008). Borrowing from the general outlines of a discourse analysis contained in Fairclough (1995), Gill (2000), and Gee (2010a, 2010b), this project first describes the context in which this language was developed; secondly, engages in textual analysis via the application of a coding scheme to “denaturalize” the text and to identify emergent patterns within the collected corpus; and lastly, reflects on the themes that emerge across the text, discussing how they relate to facets of the transmission flow framework introduced in Chapter 4.

The textual analysis in the second step of this discourse analysis involved the application of a coding tool to the text corpus. The coding tool that this study uses facilitates: first, identifying the presence of a description of an information flow within a segment of text; second, classifying the details of the information flow present in the segment according to an *a priori* schema based on the work in Chapter 4; and then finally, explicating the prevalent characteristics within the segment. The application of the coding tool is described in detail in the *Procedures* section.

Despite the use of this coding tool, this analysis still relies on a subjectively developed coding mechanism and the judgment of the researcher, and therefore is subject to criticisms and concerns around reliability and validity. In order to increase the

transparency and validity of the coding process, after the coding process, I critically examined the regularity and variability of the results, and then identified the emergent themes or notable absences in the descriptions of information flows. To further increase the reliability and validity of this work, a number of deviant cases are given particular attention in the discussion of findings.

Procedures.

As introduced in the *Instruments and measurements* section, this project relies on a three-tier coding scheme as its instrument for the textual analysis. The first stage of the coding schema required that materials be read and any descriptions or depictions of information flows be identified. As Chapter 4 argues, a transmission model of information flows considers the flows as constituted by a sender, a piece of information, a channel, and a potential receiver. However, Chapter 4 also argues that this transmission model can be expanded through the application of van Dijck's (2013) analytical framework for examining social media. Therefore, during the first stage of this coding process, each segment of text was explored for whether it mentioned: data/metadata, algorithms, protocols, interfaces, defaults, users (as senders or receivers), types of informational content on Twitter, business models, ownership status, and/or governance.

Once the presence of a description of data/metadata, algorithms, protocols, interfaces, defaults, users (as senders or receivers), types of informational content on Twitter, business models, ownership status, and/or governance was identified, the analysis moved to the second step of coding. The second step of coding involved breaking down and classifying each mention of an informational flow along the lines of whether the segment discusses *who* information flows to, *what* information is flowing,

when the information is flowing, *where* the information is flowing, *how* the information is flowing, or *why* the information is flowing. After this second tier of coding and categorization was complete, within each second tier code, a third tier of coding took place based on an inductively generated coding scheme. In this third-stage, the results of the second-stage coding were analyzed to naturally discover recurring descriptions within the secondary tier classifications (Stemler, 2001).

To illustrate how the coding process worked, it will be demonstrated using the hypothetical phrase, "Tweets are sent to the Library of Congress." This sentence can be coded on the first level as involving an information flow related to *informational content* and *users*. On the second level, as hitting the "who" (sender – Twitter [implied through passive voice]; receiver - Library of Congress) and "what" categories (Tweets). On the third-level, this analysis would look for recurring descriptions of Tweets being sent by Twitter, or the Library of Congress acting as a recipient. Dominant themes in Twitter's language use can thus be rendered visible through this data analysis process. I relied on the qualitative analysis software nVivo to facilitate the application of the coding scheme to the text corpus.

Data analysis.

Through textual analysis, this study inductively analyzes how the language Twitter, Inc. presents to users depicts informational flows on the platform. To report the outcomes of the coding process, this dissertation provides an accounting of the ways that informational flows are described within the corpus as part of Chapter 7. This reporting focuses primarily on the prevalent descriptions within each of the second level classifications; for example, highlighting particularly favored descriptions of *what* data is

made available to *whom* within the corpus. Within these second-level descriptions, this report also addresses how the facets of data/metadata, algorithms, protocols, interfaces, defaults, users, informational content, business models, ownership status or governance are described within the text corpus. Thus, this study highlights the kinds of *principles-knowledge* about Twitter's information flows which appear to be promoted within the selected language.

However, the analysis does not end at simply reporting what depictions of information flows are present. Gill (2000) notes that when doing a discourse analysis, it is important to also observe the types of descriptions and language that are not present within a text. As the Chapter 4 established what the current 'reality' of information flows is within Twitter, this discourse analysis also makes particular note of information flows that occur on Twitter, but are not present in Twitter, Inc.'s language. Through such an analysis, this work addresses the descriptions that Twitter, Inc. intentionally or unintentionally avoids.

Delimitations and limitations.

Despite best efforts to give this analysis as much rigor and structure as is reasonable, there are some delimitations and natural limitations associated with this study. First, discourse analysis has a number of natural limitations. It must be considered as a subjective interpretation generated by a researcher that exists embedded within their own subjective position (Powers, 2001). Next, the findings from this discourse analysis are not generalizable to other discourses, or even to other texts that have not been sampled as part of the corpus. While this study will make claims about the language presented to users as part of the sign-up process, these findings cannot be generalized to

things such as commercials for Twitter, other language use present on Twitter.com, or interviews that Twitter's founders may have given.

As already mentioned, discourse analysis is a subjective and inductive approach. While I have made efforts to give the findings rigor by discussing deviant cases and describing the coding process itself, it would nonetheless be possible to code and interpret this data in many different ways. This is a natural limitation of discourse analysis and should not be seen as a fatal flaw in the project. Instead, what this dissertation contributes is a unique application of discourse analysis and this particular coding frame to produce a timely and much needed analysis of Twitter, Inc.'s language.

Conclusion

By answering each of the respective operationalized research questions through the plans of study described in the previous pages, the next two chapters contribute to addressing the prime research question: In the user-Twitter relationship, what is the state of Twitter users' informational power in regards to the informational flows of the platform? The findings from each operationalized research question naturally lead into a discussion about implications for individual informational power within the context of Twitter taken up in Chapter 8. It is through the aggregation of these findings that this research makes a unique contribution of new knowledge to the current body of research on Twitter.

Chapter 6: Understandings of Information Flows on Twitter

Introduction

This chapter addresses the first research question of this project (RQ1): What knowledge of information flows do users have? As detailed Chapter 5, this project relies on an exploratory quantitative analysis of data collected via a web-based self-administered user survey distributed to over 15,000 students, faculty, and staff from a large public, urban Midwestern university. As exploratory analysis does not rely on formulating specific testable hypothesis *a priori*, this analysis tends towards a more descriptive account of the data with supplementary use of inferential statistical tools. Through exploratory analysis, this chapter draws a set of conclusions about the state of user-knowledge regarding the constitutive elements of information flows on Twitter among the sampled population.

The survey instrument prompted respondents with questions about their demographic characteristics, such as age, gender, education; about their use history with Twitter; asked a series of technical questions about information flows on the Twitter platform that were developed from the descriptive analysis of information flows on Twitter via the application of van Dijck's framework from Chapter 4; asked about their habits of consuming news about Twitter; and asked about how closely they have read the Terms of Service, Privacy Policy, and the Twitter Rules. In the context of Twitter, the respondents' understandings of the elements of information flows of the platform contribute towards the individual's *principles-knowledge*, and thus to their informational power relative to the technology. Through the analysis of the survey results, this chapter addresses RQ1, thereby generating a picture of the facets of information flow on Twitter

users have knowledge of and which they do not, thus simultaneously providing an inroad for addressing the overarching question, RQ3: "In the user-Twitter relationship, what is the state of Twitter users' informational power in regards to the informational flows of the platform?"

Following this introduction, this chapter begins by providing information about the demographic characteristics of the sample. Next, the chapter describes Twitter use patterns within the sample; provides a description of the sample's media consumption about Twitter and the sample's frequency of reading the Terms of Service, Privacy Policies, and the Twitter Rules follows. The remainder of the chapter then addresses the sample's responses to questions about some of the constitutive elements of information flows on Twitter that stem from the application of van Dijck's framework. As a side-note, the presentation of results in this chapter does not directly match the order in which questions were presented in the survey. Some of the questions in the survey were ordered in such a way not to "give away" answers to questions that came later. To help provide context for where in the survey a question appeared, question numbers have been provided in the titles of tables found in this chapter. Further, a copy of the survey instrument can be found in Appendix C.

Demographic Characteristics of the Sample Population

A total of 15,000 members of the large public, urban Midwestern University were contacted via e-mail and asked to participate in the survey in October 2014. This pool included students, faculty, and staff. According to the IT staff that facilitated the pull of the random e-mail addresses, there are roughly 60,000 active e-mail addresses at the university. Four hundred and forty-nine members of the university completed the survey,

yielding slightly less than a 3% response rate. While this fell short of the 5% response rate that was hoped for, these results may not be entirely unexpected given a lack of incentive for completion of the survey.

As described and justified in detail in the methodology chapter (Chapter 5), an attentiveness question was included in the latter third of the survey (Q35). This question stated: "The purpose of this question is to assess your attentiveness to question wording. For this question please mark the 'I don't know the answer to this question' response." Five respondents marked answers other than "I don't know the answer to this question" and were eliminated from the response pool. Another 10 respondents did not answer the question (respondents were allowed to skip questions) and were also eliminated from the pool. Thus, the final count of responses included for analysis totals 434, (N = 434).

User status.

As can be seen in the e-mail recruitment document in Appendix D, potential respondents were told, "The survey is open to all and you do not need to be a user of Twitter in order to participate." The reason the survey was made open for both, despite the fact that this project has a more concentrated interest in user understandings of information flows than non-users, is that non-users provide a useful comparison case. For example, saying that 20% of the sampled registered Twitter users can correctly identify a statement about a particular information flow as false could be an important finding in its own right, but that finding becomes much more nuanced if 20% of non-users can also identify said statement as false. If users "score" no better than non-users, careful attention should be given to why this might be the case. For example, one might ask if there is something different or lacking in the language that Twitter, Inc. uses to describe that

particular information flow to users, or in the way that particular flow is made visible/invisible to users. Inferential statistical tools, such as chi-squared analysis, can facilitate the comparison in answers among these groups to suggest whether the relationship between user status and response is statistically significant or is attributable to randomness.

As discussed in Chapter 2 however, the binary distinction between "non-users" and "users" can overlook nuances associated with different *kinds* of use and non-use. As a result, this survey uses a four-tier stratification of respondents, classifying them by whether they, "Have never used Twitter" (and are thus non-users), "Have been to the Twitter.com website, but do not have a registered account" (and are thus unregistered users), "Have a registered account on Twitter" (and are thus registered users), or "Have previously had an account on Twitter but deleted it" (and thus are formerly registered non-users). As seen in Table 6, of the 434 respondents, 25.3% indicated they are non-users, 14.3% indicated they have used the Twitter website, but do not have a registered account, 54.1% are registered users, and 6.2% are formerly registered non-users.

Table 6

Respondent User-type (Q5).

Status	Count	%
Non-user	110	25.3
Unregistered user	62	14.3
Registered user	235	54.1
Formerly registered non-user	27	6.2
Totals	434	

Note: Percentages may not total to 100% due to rounding.

This sample has a higher concentration of Twitter users than the concentration found in the U.S. broadly. Pew Research's Internet and American Life Project indicates that as of 2014, 23% of online adults use Twitter (Duggan et al., 2015). The higher

concentration of Twitter users found in this dissertation's survey data is likely attributable to multiple factors. First, the population of the university naturally does not match the overall demographic characteristics of the U.S. and thus a one to one mapping of trends between both is unlikely. Second, the recruitment method may be biased towards recruiting Twitter users. Pew's recruitment method involved random digital dialing to obtain a random sample of the individuals within the U.S. The sample in this dissertation is drawn *purposively* from a population, though through a random e-mail address selection process. Further, in the recruitment e-mail, respondents were told that the survey was interested in perceptions of Twitter, and the recruitment e-mail was titled "Research Study on Perceptions of Twitter." While individuals who opened the e-mail were told that the survey was open to both users and non-users, individuals who already have used Twitter may have been more apt to open the email and respond than non-users.

Age.

Respondents were asked three questions relating to their demographic characteristics: their age, gender, and highest level of education completed. For age, respondents were offered six potential choices: "18-24," "25-29," "30-39," "40-49," "50-59," or "60 or above." Table 7 provides a breakdown of the age responses by user-type. A total of 432 respondents answered this question, with two respondents abstaining.

Table 7

Crosstabulation of User-type by Age (Q2).

Status	18 - 24	25 - 29	30 - 39	40 - 49	50 - 59	60 or above	Totals
Non-user	24	17	18	20	11	19	109
% within group	22.0%	15.6%	16.5%	18.3%	10.1%	17.4%	
Adjusted std. residual	(-4.5)	(-.2)	(-.3)	(+2.0)	(+1.4)	(+5.8)	
Unregistered user	20	13	16	7	5	1	62
% within group	32.3%	21.0%	25.8%	11.3%	8.1%	1.6%	
Adjusted std. residual	(-1.4)	(+1.1)	(+1.8)	(-.4)	(+.3)	(-1.6)	
Registered user	114	33	41	26	15	5	234
% within group	48.7%	14.1%	17.5%	11.1%	6.4%	2.1%	
Adjusted std. residual	(+3.9)	(-1.3)	(+0)	(-1.1)	(-.7)	(-3.7)	
Formerly registered non-user	16	7	1	2	0	1	27
% within group	59.3%	25.9%	3.7%	7.4%	0.0%	3.7%	
Adjusted std. residual	(+2.1)	(+1.4)	(-2.0)	(-.9)	(-1.5)	(-.5)	
Totals	174	70	76	55	31	26	432

Note. Adjusted standardized residual frequencies appear in parentheses below observed counts.

A small number of observations can be made from Table 7. Of the respondents who answered this question (n = 432), 40.3% indicated they are 18 – 24, 16.2% indicated they are 25 - 29, 17.6% indicated they are 30 – 39, 12.7% indicated they are 40 – 49, 7.2% indicated they are 50 – 59, and 6.0% indicated they are 60 or above. Unregistered users, registered users, and formerly registered non-users tend towards a higher concentration of younger respondents than non-users, which skews slightly older. Formerly registered non-users have the highest concentration of respondents in 18 – 24 age range at 59.3%, registered users have the second highest at 48.7%, unregistered users are the third highest at 32.3%, and only 22.0% of the non-users are 18 - 24. A chi-squared test was used to explore whether there is a relationship between age and user-type within the sample. The test revealed a statistically significant relationship between whether a respondent indicates they are a "non-user," an "unregistered user," a "registered user," or a "formerly registered non-user" and their age bracket response, $X^2(15, N = 432) =$

64.145, p < .05. These observations roughly follow findings of the Pew Research's Internet Project (2014) which found individuals in the 18 – 29 year-old age range have the highest rates of being Twitter users.

Gender.

Respondents were given an open text box to indicate their gender. Four hundred thirty-one respondents gave an indication about gender, with three respondents abstaining. Responses were recoded based on the emergent categories. Table 8 provides the counts of gender by user-type.

Table 8

Crosstabulation of User-type by Gender (Q3)

User-type	Female	Genderqueer	Intersex	Male	None	Totals
Non-user	80	0	1	25	2	108
% within group	74.1%	0.0%	0.9%	23.1%	1.9%	
Adjusted std. residual	(+ 1.2)	(- .8)	(+ 1.7)	(- 2.0)	(+ 1.7)	
Unregistered user	39	0	0	23	0	62
% within group	62.9%	0.0%	0.0%	37.1%	0.0%	
Adjusted std. residual	(-.9)	(-.6)	(- . 4)	(+ 1.3)	(- .4)	
Registered user	158	1	0	75	0	234
% within group	67.5%	0.4%	0.0%	32.1%	0.0%	
Adjusted std. residual	(-.4)	(-.1)	(- 1.1)	(+ .9)	(- 1.1)	
Formerly registered non-user	18	1	0	8	0	27
% within group	66.7%	3.7%	0.0%	29.6%	0.0%	
Adjusted std. residual	(-.2)	(+ 2.6)	(-.3)	(- .1)	(- .3)	
Totals	295	2	1	131	2	431

Note: The category of "none" is not equivalent to no response. Two respondents listed "none" as a response.

Of those who gave a response to this question (n = 431), 68.4% indicated they are female, 0.4% indicated they are genderqueer, 0.2% indicated they are intersex, 30.4% indicated they are male, and 0.4% indicated they are none. This distribution is fairly even across non-users, unregistered users, registered users, and formerly registered non-users, and a subsequent chi-squared test found no statistically significant relationship between gender and user-type within the sample, $X^2(18, N = 431) = 22.543$, p = .209. Within the

group of registered Twitter users (n = 234), 67.5% indicated they are female, 0.4% indicated they are genderqueer, and 32.1% indicated they are male. This distribution includes a higher concentration of females than in the research done by the Pew Internet and American Life Project, which observed that U.S. Twitter users only slightly trend towards majority female (Brenner & Smith, 2013).

Education.

Respondents were given six possible selections from which to indicate their highest level of education completed: "Finished high school degree," "Some undergraduate education," "Finished undergraduate degree," "Some graduate-level education," "Finished graduate or other post-undergraduate professional degree." Four hundred thirty-four respondents gave an indication of their highest level of education completed. Table 9 provides the counts of highest level of education by user-type.

Table 9

Crosstabulation of User-type by Education (Q4)

Status	High School Degree	Some Undergraduate	Finished Undergraduate Degree	Some Graduate Level	Finished Graduate	Totals
Non-user	3	32	11	22	42	110
% within group	2.7%	29.1%	10.0%	20.0%	38.2%	
Adjusted std. residual	(-1.7)	(-2.0)	(-.6)	(+ 1.6)	(+ 2.1)	
Unregistered user	4	15	8	13	22	62
% within group	6.5%	24.2%	12.9%	21.0%	35.5%	
Adjusted std. residual	(+ .2)	(- 2.3)	(+ .4)	(+ 1.4)	(+ 1.0)	
Registered user	18	98	30	26	63	235
% within group	7.7%	41.7%	12.8%	11.1%	26.8%	
Adjusted std. residual	(+ 1.6)	(+ 2.2)	(+ .9)	(- 2.6)	(- 1.7)	
Formerly registered non-user	1	16	1	5	4	27
% within group	3.7%	59.3%	3.7%	18.5%	14.8%	
Adjusted std. residual	(- .5)	(+ 2.5)	(- 1.3)	(+ .5)	(- 1.8)	
Totals	26	161	50	66	131	434

Of all respondents who answered this question (n = 434), 6.0% indicated that their highest level of education completed is “Finished high school degree,” 37.1% indicated “Some undergraduate education,” 11.5% indicated “Finished undergraduate degree,” 15.2% indicated “Some graduate-level education,” and 30.2% indicated they had “Finished graduate or other post-undergraduate professional degree.” As this sample was drawn from a PhD granting institution and includes students, faculty, and staff, the distribution of educational levels does seem to make intuitive sense, although the percentage of respondents who have finished graduate degrees is far higher than the

general U.S. population.[27] Within the registered users category, a majority (41.7%) indicated "some undergraduate" as the highest level of education they have completed. This is likely attributable to the fact that, as was seen in the discussion of age, registered users in this sample tended to be younger. Conversely, among the non-user group, the highest concentration of responses was that of "finished a graduate degree" (38.2%). A chi-squared test was used to explore whether there is a significant relationship between highest level of education completed and user-type within the sample. The test revealed a statistically significant relationship between user status and what highest reported level of education completed, $X^2(12, N = 434) = 26.749$, $p < .05$. These statistics are perhaps not surprising given that there frequently is a significant relationship between age and highest level of education completed (as there was in this sample as well, as a chi-squared test confirmed, $X^2[20, N = 432] = 267.986$, $p < .05$).

Use Patterns

As technology adoption and frequency of use are not inherently uniform across all "users," these patterns can be an important intervening variable when exploring knowledge about a technology. In order to gain insight into the use patterns of unregistered users, registered users, and formerly registered non-users, the survey instrument asked a series of questions about their use behaviors in relation to the Twitter platform. These included questions about the respondents' preferred methods of accessing Twitter, how long it has been since they last visited the Twitter.com homepage, if and how long ago the respondent had last sent a "Tweet," how often the respondent accesses Twitter, and how long ago the respondent first signed up for Twitter.

[27] According to the U.S. Census Bureau (2013), the percentage of adults in the U.S. with graduate degrees is 11.6%.

Methods of access.

Respondents who have a registered account on Twitter were asked two questions about how they access Twitter. The first question (Q5a) was "How do you access Twitter? (Please select all that apply)" Respondents were given the options: "I use the Twitter.com website via my web-browser to access Twitter," "I use a desktop application, such as TweetDeck, to access Twitter," "I use a mobile application, such as the Twitter app, to access Twitter," and "I use some other means to access Twitter (please specify)" with an open text field. Of the 235 respondents with registered accounts (n = 235), 145 indicated they use the Twitter.com website (61.7% of all registered users), 19 indicated they use a desktop application (8.1% of all registered users), 191 indicated they use a mobile application (81.3% of all registered users), and 5 indicated they use some other means to access Twitter (2.1% of all registered users).[28]

Immediately following this first question, respondents were asked which way of accessing Twitter they rely on most often (Q5b). For this question, respondents could only make one selection from the same list of choices found in the previous question. Of the 235 respondents with registered accounts, 233 answered this question (n = 233). Of the 233, 57 indicated they use the Twitter.com website most often to access Twitter (24.5%), 5 indicated they use a desktop application, such as TweetDeck, most often to access Twitter (2.1%), 169 indicated they use a mobile application, such as the Twitter app, most often to access Twitter (72.5%), and 2 indicated they use some other means to access Twitter most often (0.9%).

In 2013, Twitter released a number of statistics about mobile Twitter use. In a study conducted by the research firm Compete, Twitter found that mobile is "often the

[28] Percentages do not total 100% as respondents could make multiple selections.

primary way people around the globe experience Twitter. Sixty percent of our 200 million active users log in via a mobile device at least once every month" (twitter.com, 2013, para. 2). This study also found:

> … that 18 to 34 year olds are 21% more likely to be logging into Twitter primarily via a mobile device. Not a big surprise since younger consumers tend to be stronger adopters of mobile in general. We found no statistically significant difference in the gender breakdown of primary mobile users of Twitter. It's a pretty even split. (twitter.com, 2013, para. 8)

Findings from the survey in this dissertation generally match up to the trends observed by Twitter's 2013 research. Overall, mobile apps are the most frequently used means by which to access Twitter within the sampled population. Subsequent chi-squared tests revealed a statistically significant relationship between age range and most frequently used means by which to access Twitter, $X^2(15, N = 232) = 34.109$, $p < .05$, and revealed no statistically significant relationship between gender and most frequently used means by which to access Twitter $X^2(9, N = 233) = 5.275$, $p = .810$.

Time since last visit.

Respondents who indicated that they are unregistered Twitter users, registered Twitter users, or previously registered non-users ($n = 324$) were asked when the last time they visited the Twitter.com homepage was (Q5d). They were given four options to choose from: "Over a year ago," "Over a month ago," "Over a week ago," or "Earlier this week." Table 10 provides a breakdown of the responses.

Table 10

Crosstabulation of User-type and Time of Last Visit to Twitter.com (Q5d)

User-type	Last visit > 1 year ago.	Last visit > 1 month ago.	Last visit > 1 week ago.	Earlier this week	Totals
Unregistered User	25	27	8	2	62
% within group	40.3%	43.5%	12.9%	3.2%	
Adjusted std. residual	(+ 2.9)	(+ 1.9)	(- .8)	(- 4.3)	
Registered User	45	71	41	77	234
% within group	19.2%	30.3%	17.5%	32.9%	
Adjusted std. residual	(- 4.5)	(- 1.7)	(+ .9)	(+ 5.7)	
Formerly registered non-user	14	9	4	0	27
% within group	51.9%	33.3%	14.8%	0.0%	
Adjusted std. residual	(+ 3.2)	(+ 0)	(- .2)	(- 3.1)	
Totals	84	107	53	79	323

There are a few observations to be made of this data. First, registered users have a higher respective rate of having visited the Twitter.com webpage within the earlier week (32.9%), than of unregistered users (3.2%) and formerly registered non-users (0.0%). Conversely, unregistered users and formerly registered users have a higher respective rate of having last visited the Twitter.com webpage over a year ago (40.3% and 51.9%) than registered users (19.2%). Second, a chi-squared test was used to explore whether or not there is a relationship between user-type and time of last visit to Twitter.com. The chi-squared test did reveal a statistically significant relationship between the two, $X^2(6, N = 323) = 44.075$, $p < .05$. As using Twitter.com is a means of interacting with the service for registered users, this finding does not inherently reveal anything overly surprising. Lastly, when considering all respondents ($n = 323$), the largest proportion of respondents (33.1%) indicated that their last visit to the Twitter.com was over a month ago.

Time since last Tweet.

Respondents who are registered users ($n = 235$) were asked when the last time was that they posted a Tweet (Q5c). Five possible choices were offered: “I have never

sent a Tweet," "The last time I sent a Tweet was over a year ago," "The last time I sent a Tweet was over a month ago," "The last time I sent a Tweet was over a week ago," or "The last time I sent a Tweet was earlier this week." The Twitter monitoring service "Twopcharts" estimated that in 2014, 44% of registered Twitter users have never sent a Tweet (Murphy, 2014). As Table 11 illustrates, the sample in this study has a much lower rate (10.6%) of having never sent a Tweet. This suggests that this sample may be more active on Twitter than the overall population of Twitter users.

Table 11

Last Time Registered User Sent a Tweet (Q5c)

Status	Count	%
Never sent a Tweet	25	10.6%
Last sent a Tweet over a year ago	40	17.0%
Last sent a Tweet over a month ago	46	19.6%
Last sent a Tweet over a week ago	41	17.4%
Last sent a Tweet earlier this week	83	35.3%
Totals	235	

The most frequently selected response among registered users as to when they last sent a Tweet was "earlier this week," at 35.3%. There is a fairly even distribution among the other selections.

Access rate.

Respondents who indicated that they are unregistered Twitter users, registered Twitter users, or previously registered non-users (n = 324) were asked whether they would say they access Twitter "Almost never," "Occasionally," "Fairly Regularly," or "Often" (Q5e). Table 12 provides a breakdown of the responses to this question by user type.

Table 12

Crosstabulation of User-type and Self-Described Rate of Access (Q5e)

User-type	Almost never	Occasionally	Fairly regularly	Often	Totals
Unregistered User	51	8	3	0	62
% within group	82.3%	12.9%	4.8%	0.0%	
Adjusted std. residual	(+ 6.1)	(- 2.4)	(- 1.9)	(- 3.8)	
Registered User	76	71	36	52	235
% within group	32.3%	30.2%	15.3%	22.1%	
Adjusted std. residual	(- 8.7)	(+ 3.7)	(+ 3.0)	(+ 4.8)	
Formerly registered non-user	26	1	0	0	27
% within group	96.3%	3.7%	0.0%	0.0%	
Adjusted std. residual	(+ 5.3)	(- 2.6)	(- 2.0)	(- 2.4)	
Totals	153	80	39	52	324

A subsequent chi-squared test revealed a statistically significant relationship between user-type and self-described rate of access, $X^2(6, N = 324) = 79.830$, $p < .05$. Again, this finding is unsurprising given that formerly registered non-users and unregistered users may have less impetus for accessing Twitter content than a registered user who may be more actively engaged in the service. Among registered users specifically, there is a higher concentration of respondents who would classify their use as "almost never" or "occasional" (31.3% and 30.2% respectively) than "fairly regular" or "often" (15.3% and 22.1% respectively), suggesting that roughly two-thirds of the sample self-categorize their use of Twitter as occasional or less, while the remaining third classifies their use as regular or more frequent.

Year first signed up as a registered user.

Respondents who indicated they have a registered account on Twitter or previously had an account on Twitter but deleted it were asked how long ago they first signed up for Twitter (Q5f). Respondents were given an open text field to enter their answers into. Responses were then recoded into the categories of "0 – 1 years ago," "1 – 2 years ago," "2 – 3 years ago," "3 – 4 years ago," "4 – 5 years ago," "5 – 6 years ago,"

"6 – 7 years ago," and, where text was input but was not intelligible, "Response does not indicate." Table 13 provides a breakdown of the responses by user type.

Table 13

Crosstabulation of User-type and Date of Registration (Q5f)

User-type	0-1 Years Ago	1-2 Years Ago	2-3 Years Ago	3-4 Years Ago	4-5 Years Ago	5-6 Years Ago	6-7 Years Ago	Response does not indicate	Total
Registered User	32	59	55	35	19	7	3	2	212
% within group	15.1%	27.8%	25.9%	16.5%	9.0%	3.3%	1.4%	0.9%	
Adjusted std. residual	(+ .4)	(+ .8)	(+ .2)	(- .4)	(+ .2)	(- 2.9)	(+ .6)	(+ .5)	
Formerly registered non-user	3	5	6	5	2	4	0	0	25
% within group	12.0%	20.0%	24.0%	20.0%	8.0%	16.0%	0.0%	0.0%	
Adjusted std. residual	(- .4)	(- .8)	(- .2)	(+ .4)	(- .2)	(+ 2.9)	(- .6)	(- .5)	
Totals	35	64	61	40	21	11	3	2	237

A chi-squared test revealed no statistically significant relationship between user-type and date of sign-up, $X^2(7, N = 237) = 9.229$, p = .237. This suggests that there is not inherently a relationship between how long ago a user signed up and whether they are now a registered user or a formerly registered non-user within the sample. Further, this suggests that deleting one's account may not be associated with length of time as a Twitter user. Of registered users that responded (n = 212), 68.8% signed up for Twitter in the past three years, whereas only 30.2% signed up more than three years ago. This suggests that registered users in the sample skew towards having spent fewer years on the service.

Media Consumption and Policy Document Reading Habits

In addition to the questions about demographic characteristics and usage habits, at the end of the survey, respondents were asked a set of questions about how often they hear Twitter discussed in the news (Q47), how often they themselves keep up with news

about Twitter (Q48), and about how closely they have read the Terms of Service (Q49), Privacy Policies (Q50), and the Twitter Rules (Q51). These questions were asked at the end of the survey in order to not "give away" other questions that appeared earlier in the survey (specifically Q32).

The first prompt participants were given as part of this set was, "Pick the option that would best describe your opinion: I feel like Twitter is discussed in the news…" Participants could then select "Never," "Sometimes," "Occasionally," or "Regularly." Table 14 provides a breakdown of the responses to this statement, stratified by user-type.

Table 14

Crosstabulation of User-type and Self-Reported Frequency of Hearing News about Twitter (Q47)

Status	Never	Occasionally	Sometimes	Regularly	Totals
Non-user	5	25	28	51	109
% within group	4.6%	22.9%	25.7%	46.8%	
Adjusted std. residual	(+ 1.8)	(- .6)	(- .3)	(+ .2)	
Unregistered user	1	15	17	29	62
% within group	1.6%	24.2%	27.4%	46.8%	
Adjusted std. residual	(- .4)	(- .2)	(+ .2)	(+ .1)	
Registered user	4	64	64	102	234
% within group	1.7%	27.4%	27.4%	43.6%	
Adjusted std. residual	(- .9)	(+ 1.2)	(+ .4)	(- 1.1)	
Formerly reg. non-user	0	4	6	17	27
% within group	0.0%	14.8%	22.2%	63.0%	
Adjusted std. residual	(- .8)	(- 1.3)	(- .5)	(+ 1.8)	
Totals	10	108	115	199	432

A majority of each group indicated that they hear Twitter discussed in the news regularly, with formerly registered non-users (n = 27) having indicated at the highest rate (63.0%) and registered users (n = 234) at the lowest rate (43.6%). Only 10 of 432 respondents (2.3%) indicated that they "never" heard Twitter discussed in the news. A chi-squared test revealed no statistically significant relationship between user-type and how often

respondents indicated they heard Twitter discussed in the news, $X^2(9, N = 432) = 7.696$, p = .565.

The second prompt participants were given as part of this set was, "Pick the option that would best describe you: I keep up with news about Twitter…" Participants could then select "Never," "Sometimes," "Occasionally," or "Regularly." Table 15 provides a breakdown of the responses to this statement, stratified by user-type.

Table 15

Crosstabulation of User-type and Self-Reported Keeping up with News about Twitter (Q48)

Status	Never	Occasionally	Sometimes	Regularly	Totals
Non-user	79	26	3	1	109
% within group	72.5%	23.9%	2.8%	0.9%	
Adjusted std. residual	(+ 5.6)	(- 2.6)	(- 3.6)	(- 1.8)	
Unregistered user	31	24	7	0	62
% within group	50.0%	38.7%	11.3%	0.0%	
Adjusted std. residual	(+ .1)	(+ .8)	(- .4)	(- 1.7)	
Registered user	88	90	42	15	235
% within group	37.4%	38.3%	17.9%	6.4%	
Adjusted std. residual	(- 5.4)	(+ 2.0)	(+ 3.5)	(+ 3.2)	
Formerly reg. non-user	16	8	3	0	27
% within group	59.3%	29.6%	11.1%	0.0%	
Adjusted std. residual	(+ 1.1)	(- .5)	(- .3)	(- 1.1)	
Totals	214	148	55	16	433

With the exception of registered users, a majority of all other user groups indicated that they "never" keep with news about Twitter. Registered users within the sample, however, indicated they keep up with news about Twitter at a slightly higher frequency. Of registered users (n = 235), 37.4% indicated never keeping up with news about Twitter, whereas 38.3% indicated they do so occasionally, 17.9% indicated they do so sometimes, and 6.4% indicated they do so regularly. However, due to the low number of overall respondents who indicated they regularly and even sometimes keep up with news about

Twitter across all groups (four cells (25.0%) have an expected count of less than 5) the conditions for the use of a chi-squared test were not met. Thus it remains unknown whether there is a statistically significant relationship between user-status and keeping up with news about Twitter.

The third prompt participants were given as part of this set was, "Users of Twitter who have registered accounts have to agree to the Terms of Service when they sign up for the service. Select the option below that best describes how closely you would say you have read the Terms of Service." Participants could then select "I have never read the Terms of Service" "I have skimmed over the Terms of Service," "I have read the Terms of Service in some detail, but not fully," or "I have fully read the Terms of Service in detail." Table 16 provides a breakdown of the responses to the question, stratified by user-type.

Table 16

Crosstabulation of User-type and Self-Reported Reading of Terms of Service (Q49)

Status	Never	Skimmed	In some detail	Fully	Totals
Non-user	107	1	1	0	109
% within group	98.2%	0.9%	0.9%	0.0%	
Adjusted std. residual	(+ 6.0)	(- 5.9)	(- .5)	(- .8)	
Unregistered user	56	6	0	0	62
% within group	90.3%	9.7%	0.0%	0.0%	
Adjusted std. residual	(+ 2.6)	(- 2.3)	(- 1.0)	(- .6)	
Registered user	155	73	5	2	235
% within group	66.0%	31.1%	2.1%	0.9%	
Adjusted std. residual	(- 6.2)	(+ 5.7)	(+ 1.4)	(+ 1.3)	
Formerly reg. non-user	17	10	0	0	27
% within group	63.0%	37.0%	0.0%	0.0%	
Adjusted std. residual	(- 1.8)	(+ 2.1)	(- .6)	(- .4)	
Totals	335	90	6	2	433

Perhaps unsurprisingly, few respondents overall have read the terms of service either fully or in some detail (only 8 of 433, or 1.8%). More than half of each group

indicated that they have never read the Terms of Service, with registered users (n = 235) having indicated "never" at a rate of 66.0%. Only 34.1% of registered users indicated that they have at least skimmed the Terms of Service. Respondents to this survey indicated a much lower rate of reading terms of service than in the study done by Fuchs (2009) which explored student knowledge of MySpace, Facebook and studiVZ users, and included questions about terms of service reading habits.

The fourth prompt participants were given as part of this set was, "Users of Twitter who have registered accounts have to agree to the Privacy Policy when they sign up for the service. Select the option below that best describes how closely you would say you have read the Privacy Policies." Participants could then select "I have never read the Privacy Policies" "I have skimmed over the Privacy Policies," "I have read the Privacy Policies in some detail, but not fully," or "I have fully read the Privacy Policies in detail." Table 17 provides a breakdown of the responses to this statement, stratified by user-type.

Table 17

Crosstabulation of User-type and Self-Reported Reading of Privacy Policies (Q50)

Status	Never	Skimmed	In some detail	Fully	Totals
Non-user	104	2	1	1	108
% within group	96.3%	1.9%	0.9%	0.9%	
Adjusted std. residual	(+ 5.2)	(- 5.1)	(- 1.0)	(+ .3)	
Unregistered user	56	5	0	1	62
% within group	90.3%	8.1%	0.0%	1.6%	
Adjusted std. residual	(+ 2.4)	(- 2.3)	(- 1.2)	(+ .9)	
Registered user	163	64	7	1	235
% within group	69.4%	27.2%	3.0%	0.4%	
Adjusted std. residual	(- 5.2)	(+ 5.1)	(+ 1.4)	(- .7)	
Formerly reg. non-user	17	9	1	0	27
% within group	63.0%	33.3%	3.7%	0.0%	
Adjusted std. residual	(- 2.1)	(+ 2.0)	(+ .6)	(- .4)	
Totals	340	80	9	3	432

Similar to the reading habits around the Terms of Service, few of the respondents have read the Privacy Policies either fully or in some detail (only 12 of 432, or 2.8%). More than half of each group indicated that they have never read the Privacy Policies, with registered users (n = 235) having indicated "never" at a rate of 69.4%. Only 30.6% of registered users indicated that they have at least skimmed the Terms of Service.

The fifth prompt participants were given as part of this set was, "Users of Twitter who have registered accounts have to agree to the Twitter Rules when they sign up for the service. Select the option below that best describes how closely you would say you have read the Twitter Rules." Participants could then select "I have never read the Twitter Rules" "I have skimmed over the Twitter Rules," "I have read the Twitter Rules in some detail, but not fully," or "I have fully read the Twitter Rules in detail." Table 18 provides a breakdown of the responses to this question, stratified by user-type.

Table 18

Crosstabulation of User-type and Self-Reported Reading of Twitter Rules (Q51)

Status	Never	Skimmed	In some detail	Fully	Totals
Non-user	106	2	1	0	109
% within group	97.2%	1.8%	0.9%	0.0%	
Adjusted std. residual	(- 3.9)	(- .4)	(+ 4.0)	(+ 4.0)	
Unregistered user	57	5	0	0	62
% within group	91.9%	8.1%	0.0%	0.0%	
Adjusted std. residual	(- .4)	(- .4)	(+ .5)	(+ .5)	
Registered user	191	35	6	2	234
% within group	81.6%	15.0%	2.6%	0.9%	
Adjusted std. residual	(+ 3.1)	(+ 1.0)	(- 3.4)	(- 3.4)	
Formerly reg. non-user	21	5	1	0	27
% within group	77.8%	18.5%	3.7%	0.0%	
Adjusted std. residual	(+ 1.2)	(- .8)	(- .9)	(- .9)	
Totals	375	47	8	2	432

Similar to the other two governing documents, few of the respondents have read the Twitter Rules either fully or in some detail (only 10 of 432, or 2.3%). More than three-quarters of each group indicated that they have never read the Twitter Rules, with registered users (n = 234) having indicated "never" at a rate of 81.6%. Only 18.4% of registered users indicated that they have at least skimmed the Twitter Rules. Among the three governing documents, the Twitter Rules are the least frequently read.

Knowledge of Data/Metadata among Sample

After the questions on use habits (where applicable), participants were presented with the first of nine sets of questions on specific components of information flows on Twitter; questions related to data and metadata on Twitter. Based on the information flow framework developed in Chapter 4 (see Figure 5), data and metadata structures are part of the techno-cultural formation that constitute part of the "channel" of information flow. For the questions in this section, respondents were instructed in the following way:

> For the following questions, we will be asking you a series of questions about Twitter. Please select what you believe to be the correct answer. Please do not use any external sources to help you answer these questions. You may choose "I don't know the answer to this question" if you feel that you do not know the answer.

Participants were then given a statement about data/metadata on Twitter that was either accurate or inaccurate, and were asked to indicate that "Yes, this [statement] is correct," "No, this is incorrect," or "I don't know the answer to this question."

The first statement that respondents were given (Q6) relates to the message length on the Twitter platform. The statement reads: "Messages on Twitter (also called 'Tweets') are limited to 210-characters in length." Messages on Twitter are actually limited to 140-characters, and hence, this statement is incorrect. Table 19 shows the results of this question, broken down by user type.

Table 19

Crosstabulation of User-type and Info Flow Question (IFQ): Tweet Length (Q6)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	27	25	58	110
% within group	24.5%	22.7%	52.7%	
Adjusted std. residual	(+ 2.2)	(- 6.7)	(+ 5.5)	
Unregistered user	12	28	22	62
% within group	19.4%	45.2%	35.5%	
Adjusted std. residual	(+ .4)	(- .9)	(+ .7)	
Registered user	34	152	49	235
% within group	14.5%	64.7%	20.9%	
Adjusted std. residual	(- 1.9)	(+ 6.4)	(- 5.3)	
Formerly reg. non-user	4	14	9	27
% within group	14.8%	51.9%	33.3%	
Adjusted std. residual	(- .4)	(+ .1)	(+ .2)	
Totals	77	219	138	434

Note. The statement for this question was incorrect.

Registered users (n = 235) had the highest rate (64.7%) of accurately identifying this statement as incorrect among all user-type groups, had the lowest rate of inaccurately identifying the statement as correct at 14.5%, and had the lowest "do not know" response rate at 20.9%. However, as Twitter provides a running count of the number of characters left for users as they compose a Tweet within the "New Tweet" interface, it was somewhat surprising to see that only two-thirds of registered users could accurately identify the original statement as incorrect. Non-users (n = 110) had both the highest rate of inaccurately identifying the statement as correct (24.5%) and had the highest rate of indicating that they did not know the answer (52.7%). A subsequent chi-squared test confirmed a statistically significant relationship between user-type and response to this question, $X^2(6, N = 434) = 55.345$, $p < .05$.

Next, respondents were given the correct statement: "Users can select an option when composing a Tweet to share location information, such as their GPS coordinates,

along with their Tweet." Table 20 shows the results of this question, broken down by user type.

Table 20

Crosstabulation of User-type and IFQ: Users Can Share GPS Data in Tweets (Q7)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	35	4	71	110
% within group	31.8%	3.6%	64.5%	
Adjusted std. residual	(- 2.5)	(- 2.0)	(+ 3.6)	
Unregistered user	16	5	40	61
% within group	26.2%	8.2%	65.6%	
Adjusted std. residual	(- 2.7)	(+ .0)	(+ 2.6)	
Registered user	119	23	91	233
% within group	51.1%	9.9%	39.1%	
Adjusted std. residual	(+ 4.1)	(+ 1.4)	(- 4.9)	
Formerly reg. non-user	11	3	13	27
% within group	40.7%	11.1%	48.1%	
Adjusted std. residual	(- .1)	(+ .6)	(- .2)	
Totals	181	35	215	431

Note. The statement for this question is correct.

Registered users (n = 233) had the highest rate of accurately identifying the statement as correct (51.1%), but also had the second highest rate of inaccurately identifying this statement as incorrect (9.9%). While registered users having the highest rate of accurately identifying this statement as correct should not come as a surprise, what is somewhat surprising is that only slightly over half of the registered users could make this accurate identification, given that the functionality is part of the Tweet creation interface. Among non-users (n = 110) and unregistered users (n = 61), there appears to be a higher degree of uncertainty about the veracity of the statement (64.5% and 65.6% respectively), whereas fewer than 40% of registered users indicated they did not know the answer to the question. Subsequent chi-squared tests confirmed a statistically significant relationship between user-type and response to this question, $X^2(6, N = 431) = 28.050$, $p < .05$.

Next, respondents were given the incorrect statement: "Twitter does not ever withhold Tweets or user accounts from being accessed within specific countries, even if they have received a legal request to do so." This statement is inaccurate as Twitter does withhold Tweets if they have received a legal request to do so (Mackey, 2014), and the technical code that facilitates implementation of this feature is part of a given Tweet's metadata. Table 21 shows the results of this question, broken down by user type.

Table 21

Crosstabulation of User-type and IFQ: Tweets Ever Withheld (Q8)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	5	18	87	110
% within group	4.5%	16.4%	79.1%	
Adjusted std. residual	(- .9)	(- .6)	(+ 1.1)	
Unregistered user	2	16	44	62
% within group	3.2%	25.8%	71.0%	
Adjusted std. residual	(-1.1)	(+ 1.7)	(- .9)	
Registered user	20	42	172	234
% within group	8.5%	17.9%	73.5%	
Adjusted std. residual	(+1.9)	(- .2)	(- .9)	
Formerly reg. non-user	1	3	23	27
% within group	3.7%	11.1%	85.2%	
Adjusted std. residual	(- .6)	(- 1.0)	(+ 1.2)	
Totals	28	79	326	433

Note. The statement for this question is incorrect.

Registered users (n = 234) inaccurately identified this statement as correct at a rate of 8.5% (the highest among all user-groups), accurately identified this statement as incorrect at a rate of 17.9%, and indicated that they did not know whether Tweets were ever withheld at a rate of 73.5%. These findings do make a certain amount of sense given that this practice is relatively new and that Twitter has blocked relatively few Tweets (at least in comparison to the overall number of Tweets that exist on Twitter), and thus, registered users are relatively unlikely to come across these Tweets as part of their

timelines. Overall, more than 70% of each user group indicated that they did not to know the answer to this question. A subsequent chi-squared test found no statistically significant relationship between user-type and response to this question, $X^2(6, N = 433) = 7.256$, $p = .298$.

Respondents were next presented with the incorrect statement: "If you have a "protected" account on Twitter, your Tweets are only visible to the users that follow you and to the users that follow your followers." While it is true that if you have a "protected" account on Twitter, your Tweets are visible to the users that follow you, it is not true that your Tweets are also visible to the users that follow your followers, hence the statement is incorrect. Table 22 provides a breakdown of the responses to this statement, stratified by user-type.

Table 22

Crosstabulation of User-type and IFQ: Protected Tweets Visible to Follower-Followers (Q9)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	25	4	81	110
% within group	22.7%	3.6%	73.6%	
Adjusted std. residual	(- 2.7)	(- 4.9)	(+ 6.5)	
Unregistered user	25	5	32	62
% within group	40.3%	8.1%	51.6%	
Adjusted std. residual	(+ 1.3)	(- 2.5)	(+ .8)	
Registered user	84	72	79	235
% within group	35.7%	30.6%	33.6%	
Adjusted std. residual	(+ 1.2)	(+ 6.1)	(- 6.1)	
Formerly reg. non-user	10	5	12	27
% within group	37.0%	18.5%	44.4%	
Adjusted std. residual	(+ .4)	(- .2)	(- .3)	
Totals	144	86	204	434

Note. The statement for this question is incorrect.

Among registered users ($n = 235$), slightly more than one-third inaccurately indicated the statement as correct (35.7%), slightly less than one-third accurately indicated the statement is incorrect (30.6%), and almost one-third indicated they do not

know (33.6%). While registered users do have the highest rate of accurately identifying the statement as incorrect among all user-types, a two-thirds majority of registered users had either inaccurate or uncertain responses about how Tweets flow to followers of followers when a user chooses to "protect" their accounts. This finding is particularly important as, outside of choosing whether or not to "protect" one's account on Twitter, there are few other privacy controls users have. Among other user-type groups, there is a much higher rate of "don't know" responses (44.4% among formerly registered non-users, 51.6% among unregistered users, and 73.6% among non-users) and a far lower rate of accurately identifying the statement as incorrect (18.5% among formerly registered non-users, 8.1% among unregistered users, and 3.6% among non-users). A subsequent chi-squared test confirmed a statistically significant relationship between user-type and response to this question, $X^2(6, N = 434) = 63.804$, $p < .05$.

The final statement respondents were given as part of this set of questions is the correct statement: "Twitter warns users if a link embedded in someone else's Tweet has been marked as 'possibly sensitive.'" Table 23 provides a breakdown of the responses to this statement, stratified by user-type.

Table 23

Crosstabulation of User-type and IFQ: Twitter Warns about Possibly Sensitive Tweets (Q10)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	9	3	97	109
% within group	8.3%	2.8%	89.0%	
Adjusted std. residual	(- 2.7)	(- 3.0)	(+ 4.3)	
Unregistered user	14	3	45	62
% within group	22.6%	4.8%	72.6%	
Adjusted std. residual	(+ 1.4)	(- 1.6)	(- .1)	
Registered user	44	36	154	234
% within group	18.8%	15.4%	65.8%	
Adjusted std. residual	(+ 1.4)	(+ 3.7)	(-3.7)	
Formerly reg. non-user	4	3	20	27
% within group	14.8%	11.1%	74.1%	
Adjusted std. residual	(- .2)	(+ .1)	(+ .1)	
Totals	71	45	316	432

Note. The statement for this question is correct.

Registered users (n = 234) had the highest rate of inaccurately identifying the statement as incorrect (15.4%), but also had the second highest rate of accurately identifying the statement as correct (18.8%). Just over 65% of registered users indicated they did not know. This suggests that the "possibly sensitive" warning mechanism is not particularly well understood by registered users. Overall, more than 65% of each user-type group responded that they did not know the veracity of this statement. A chi-squared test revealed a statistically significant relationship between user-type and response to this question, $X^2(6, N = 432) = 25.725$, $p < .05$.

Knowledge of Protocols among Sample

Following the questions on data/metadata, participants were brought to the second of nine sets of questions regarding information flows on Twitter, questions about the protocols on the site. Based on the information flow framework developed in Chapter 4 (see Figure 5), protocols are part of the techno-cultural formation that constitute part of the "channel" of information flow. For this set, respondents were again instructed that they would be given a correct or incorrect statement and would then be asked to indicate

that “Yes, this [statement] is correct,” “No, this is incorrect,” or “I don’t know the answer to this question.”

The first statement participants were given as part of this set is correct statement: “Including a hashtag (the ‘#’ symbol) in front of a word is a way of marking keywords or topics in a Tweet and is sometimes used by users as a way to categorize messages.” Table 24 provides a breakdown of the responses to this statement, stratified by user-type.

Table 24

Crosstabulation of User-type and IFQ: How Hashtags Work (Q11)

Status	Statement is Correct	Statement is Incorrect	Don’t Know	Totals
Non-user	78	4	28	110
% within group	70.9%	3.6%	25.5%	
Adjusted std. residual	(- 7.0)	(+ 1.6)	(+ 6.8)	
Unregistered user	59	0	3	62
% within group	95.2%	0.0%	4.8%	
Adjusted std. residual	(+ 1.7)	(- 1.2)	(- 1.3)	
Registered user	223	4	8	235
% within group	94.9%	1.7%	3.4%	
Adjusted std. residual	(+ 4.3)	(- .2)	(- 4.5)	
Formerly reg. non-user	26	0	1	27
% within group	96.3%	0.0%	3.7%	
Adjusted std. residual	(+ 1.3)	(- .7)	(- 1.0)	
Totals	386	8	40	434

Note. The statement for this question is correct.

The clear majority of registered users (*n* = 235) were able to accurately identify this statement as correct (94.9%). Only 1.7% inaccurately identified this statement as incorrect, and only 3.4% indicated they did not know. This demonstrates high familiarity among registered users of the protocological functions of hashtags. This familiarity also extended to other users groups: over 90% of unregistered users (*n* = 62) and formerly registered non-users (*n* = 27) accurately identified this statement as correct. Even among non-users (*n* = 110), the protocological function of hashtags appears to be fairly well understood, with over 70% accurately identifying the statement as correct. In fact, across all user type groups, only eight respondents inaccurately identified the statement as

incorrect. As hashtags are regularly appearing on commercials and on other social media platforms and appear to have entered into popular vernacular, this high level of familiarity perhaps should not come as a surprise. As a result of violations to the assumptions that undergird the use of chi-squared tests for this question (as five cells (41.7%) had expected count less than five) a chi-squared test could not be used to test whether a significant relationship exists between the two categories.

The next statement focused on how @replies function. Respondents were given the accurate statement: "Including the "@" symbol and another user's Twitter username (such as "@PBS") at the beginning of a Tweet means that Twitter will treat that as a "reply" to that user." Table 25 provides a breakdown of the responses to this statement, stratified by user-type.

Table 25

Crosstabulation of User-type and IFQ: What Makes a Tweet a @Reply (Q12)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	36	6	68	110
% within group	32.7%	5.5%	61.8%	
Adjusted std. residual	(- 7.5)	(- .9)	(+ 8.4)	
Unregistered user	42	0	20	62
% within group	67.7%	0.0%	32.3%	
Adjusted std. residual	(+ .9)	(- 2.4)	(+ .4)	
Registered user	170	26	39	235
% within group	72.3%	11.1%	16.6%	
Adjusted std. residual	(+ 4.5)	(+ 3.2)	(- 6.6)	
Formerly reg. non-user	24	0	3	27
% within group	88.9%	0.0%	11.1%	
Adjusted std. residual	(+ 2.9)	(- 1.5)	(- 2.2)	
Totals	272	32	130	434

Note. The statement for this question is correct.

While registered users (n = 235) were less certain about this statement than the statement about hashtags, almost three-quarters of registered users were able to accurately identify

the statement as correct. However, 11.1% of registered users inaccurately identified this statement as incorrect, and 16.6% indicated that they do not know if the statement is correct or incorrect. Among non-users (n = 110), there is a much lower percentage (32.7%) of accurate identification of the statement as correct. This suggests that @replies have not trickled into the non-user consciousness to the same degree that hashtags have. A follow-up chi-squared analysis revealed a statistically significant relationship between user-status and response to this question, $X^2(6, N = 434) = 88.507$, $p < .05$.

Next, respondents were presented with the incorrect statement: "If you "favorite" another user's Tweet on Twitter, that information is only shared with the person who created the Tweet." This statement is incorrect because a user's "favorites" are publicly accessible for "public" users, and accessible by a "protected" user's followers. Table 26 provides a breakdown of the responses to this statement, stratified by user-type.

Table 26

Crosstabulation of User-type and IFQ: Favorites Accessible to Others (Besides Author) (Q13)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	1	15	92	108
% within group	0.9%	13.9%	85.2%	
Adjusted std. residual	(- 3.3)	(- 7.1)	(+ 8.9)	
Unregistered user	3	26	33	62
% within group	4.8%	41.9%	53.2%	
Adjusted std. residual	(- 1.1)	(-.2)	(+ .8)	
Registered user	31	136	67	234
% within group	13.2%	58.1%	28.6%	
Adjusted std. residual	(+ 3.8)	(+ 6.8)	(- 8.9)	
Formerly reg. non-user	2	9	16	27
% within group	7.4%	33.3%	59.3%	
Adjusted std. residual	(- .2)	(- 1.1)	(+ 1.2)	
Totals	37	186	208	431

Note. The statement for this question is incorrect.

Less than sixty percent of registered users (n = 234) accurately identified this statement as incorrect (58.1%), 13.2% inaccurately identified this statement as correct, and 28.6% indicated that they did not know. Registered users had the highest rate of inaccurately identifying the statement as correct across all user-types. This suggests that while a majority of registered users understand the accessibility of information generated through the favorites protocol, there appears to be some uncertain and incorrect knowledge among about 40% of the sampled registered users. Among all other groups, the rate of accurately identifying the statement as incorrect fell below 50%, and in the case of non-users (n = 108), fell rather dramatically to 13.9%. The "don't know" response rate also grew, with 85.2% non-users, 59.3% of formerly registered non-users, and 53.2% of unregistered users indicating they did not know whether the statement was correct or incorrect. A chi-squared analysis revealed a statistically significant relationship between user-status and response to this question, $X^2(6, N = 431) = 98.752$, $p < .05$.

The last statement respondents were given as part of this set is the correct statement: "Following someone on Twitter means that you are subscribing to their Tweets as a follower and their updates will appear in your Home tab." Table 27 provides a breakdown of the responses to this statement, stratified by user-type.

Table 27

Crosstabulation of User-type and IFQ: How Following Works (Q14)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	66	0	44	110
% within group	60.0%	0.0%	40.0%	
Adjusted std. residual	(- 7.2)	(- .8)	(+ 7.4)	
Unregistered user	49	0	13	62
% within group	79.0%	0.0%	21.0%	
Adjusted std. residual	(- .8)	(- .6)	(+ .9)	
Registered user	217	2	15	234
% within group	92.7%	0.9%	6.4%	
Adjusted std. residual	(+ 6.1)	(+ 1.3)	(- 6.4)	
Formerly reg. non-user	25	0	2	27
% within group	92.6%	0.0%	7.4%	
Adjusted std. residual	(+ 1.4)	(- .4)	(- 1.4)	
Totals	357	2	74	433

Note. The statement for this question is correct.

Registered users (n = 234) had the highest rate of accurate identification of the statement as incorrect among all user-types at 92.7%. Only 2 of 234 registered users inaccurately identified this statement as incorrect (0.9%) and only 15 of 234 indicated they did not know (6.4%). This shows a relatively high understanding of the following protocol among registered users overall. More than half of the respondents across all user-type groups could accurately identify this statement as correct, and across all groups, there were almost no inaccurate identifications of this statement as incorrect. However, the non-user group did have a rate of indicating that they "don't know" of 40.0%. A chi-squared analysis revealed a statistically significant relationship between user-status and response to this question, $X^2(6, N = 433) = 63.279$, $p < .05$.

Knowledge of Algorithms among Sample

Following the questions on protocols, participants were brought to the third of nine sets of questions regarding information flows on Twitter, questions about the

algorithms on the site. Based on the information flow framework developed in Chapter 4 (see Figure 5), algorithms structures are part of the techno-cultural formation that constitute part of the "channel" of information flow. For this set, respondents were again instructed that they would be given a correct or incorrect statement and would then be asked to indicate that "Yes, this [statement] is correct," "No, this is incorrect," or "I don't know the answer to this question."

The first statement that respondents were presented with as part of this series is the incorrect statement: "Twitter user "@Jane" has a protected account. @Jane sends an @reply to Twitter user "@PBS." @PBS will still be able to see @Jane's Tweet, even if they are not following her." This statement is incorrect because, if they are not following @Jane, @PBS will not be able to see @Jane's "protected" Tweets, even if they include mention of @PBS. Table 28 provides a breakdown of the responses to this statement, stratified by user-type.

Table 28

Crosstabulation of User-type and IFQ: Protected @Replies (Q15)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	22	1	87	110
% within group	20.0%	0.9%	79.1%	
Adjusted std. residual	(- 5.1)	(- 2.5)	(+ 6.2)	
Unregistered user	25	2	35	62
% within group	40.3%	3.2%	56.5%	
Adjusted std. residual	(- .1)	(- .9)	(+ .5)	
Registered user	121	20	94	235
% within group	51.5%	8.5%	40.0%	
Adjusted std. residual	(+ 4.9)	(+ 3.0)	(- 6.2)	
Formerly reg. non-user	9	1	17	27
% within group	33.3%	3.7%	63.0%	
Adjusted std. residual	(- .8)	(- .4)	(+ 1.0)	
Totals	177	24	233	434

Note. The statement for this question is incorrect.

More than half of registered users (n = 235) inaccurately identified this as a correct statement (51.5%). While registered users also have the highest rate of having accurately identified this statement as incorrect, they did so only at a rate of 8.5%. Exactly 40% of registered users indicated that they did not know whether this statement is correct. This suggests that how @replies flow when a user has a protected account may be poorly understood by registered users. Meanwhile, more than half of the other user-type groups indicated they did not know whether the statement is correct or incorrect. A subsequent chi-squared test showed a statistically significant relationship between user-type and response to this question, $X^2(6, N = 434) = 49.351$, $p < .05$. Overall, the data suggests that the ways Tweets flow from protected accounts when those Tweets invoke @reply algorithms is not well understood by all user-type groups, but may be particular *misunderstood* by registered users.

The next statement users were given is the correct statement, "Twitter's trending topic algorithm identifies topics that are immediately popular, rather than topics that have been popular for a while or on a daily basis." The issue of when and how trending topics become visible was particularly salient during the Occupy Wall Street protests. Some protestors accused Twitter of censorship when occupy related hashtags failed to show up as part of the "trending topics" despite their popularity in use. Some outside commentators observed that this may not be active censorship, but simply an artifact of how Twitter's algorithms determine what shows up as a trending topic, and suggested that critics may not be fully aware of how Twitter's algorithms function (RT News, 2011). Table 29 provides a breakdown of the responses to this statement, stratified by user-type.

Table 29

Crosstabulation of User-type and IFQ: Immediate Popularity of Trending Topics (Q16)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	25	2	82	109
% within group	22.9%	1.8%	75.2%	
Adjusted std. residual	(- 4.4)	(- 1.9)	(+ 5.1)	
Unregistered user	19	4	39	62
% within group	30.6%	6.5%	62.9%	
Adjusted std. residual	(- 1.7)	(+ .4)	(+ 1.5)	
Registered user	122	15	98	235
% within group	51.9%	6.4%	41.7%	
Adjusted std. residual	(+ 5.2)	(+ 1.1)	(- 5.6)	
Formerly reg. non-user	10	2	15	27
% within group	37.0%	7.4%	55.6%	
Adjusted std. residual	(- .4)	(+ .5)	(+ .2)	
Totals	176	23	234	433

Note. The statement for this question is correct.

More than half of registered users (n = 235) were able to accurately identify this statement as correct (51.9%). However, 41.7% indicated that they did not know if the statement was correct or incorrect, and 6.4% inaccurately identified the statement as incorrect. This suggests that, while a slim majority of respondents who are registered users could identify this statement as accurate, many were uncertain about this particular mechanic of the trending topics algorithm. Across all groups, only 23 of 433 respondents inaccurately identified the answer to this question as incorrect. However, particularly among non-users, unregistered users, and formerly registered non-users, the "don't know" response was chosen more than half the time (75.2% for non-users, 62.9% for unregistered users, and 55.6% for formerly registered non-users). A subsequent chi-squared test showed a statistically significant relationship between user-type and response choice, $X^2(6, N = 433) = 37.316$, $p < .05$.

Staying with the trending topic algorithm, respondents were next presented with the incorrect statement: "All users see the same trending topics." This statement is incorrect as all users do not see the same trending topics. Table 30 provides a breakdown of the responses to this statement, stratified by user-type.

Table 30

Crosstabulation of User-type and IFQ: What Trending Topics Users See (Q17)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	7	19	83	108
% within group	6.5%	17.6%	76.9%	
Adjusted std. residual	(- 2.0)	(- 4.8)	(+ 5.9)	
Unregistered user	4	26	32	62
% within group	6.5%	41.9%	51.6%	
Adjusted std. residual	(- 1.4)	(+ .9)	(+ .0)	
Registered user	34	106	93	234
% within group	14.5%	45.3%	39.7%	
Adjusted std. residual	(+ 2.1)	(+ 4.1)	(- 5.3)	
Formerly reg. non-user	5	7	15	27
% within group	18.5%	25.9%	55.6%	
Adjusted std. residual	(+ 1.2)	(- 1.2)	(+ .4)	
Totals	37	186	208	431

Note. The statement for this question is incorrect.

Registered users (n = 234) were able to accurately identify this statement as incorrect at a rate of 45.3%, the highest among all user-types. However, registered users also had the second highest rate of inaccurately identifying the statement as incorrect at a rate of 14.5%, and 39.7% of registered users indicated they were not sure if this statement was correct or incorrect. This suggests that what others see via the trending topics algorithm may not be well understood by registered users. Among non-users, unregistered users, and formerly registered non-users, more than half in each group indicated that they did not know the veracity of the statement, with non-users doing so at a rate of 76.9%, unregistered users at a rate of 51.6%, and formerly registered non-users at a rate of

55.6%. A subsequent chi-squared test showed a statistically significant relationship between user-type and response choice, $X^2(6, N = 431) = 43.044$, $p < .05$.

Next, respondents were given the correct statement: "A Twitter user will only see another user's @replies in their home timeline if they are following both the sender and recipient of the @reply." Table 31 provides a breakdown of the responses to this statement, stratified by user-type.

Table 31

Crosstabulation of User-type and IFQ: @Replies Visibility on Followers' Timelines (Q18)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	6	11	93	110
% within group	5.5%	10.0%	84.5%	
Adjusted std. residual	(- 3.6)	(- 2.9)	(+ 5.2)	
Unregistered user	5	9	47	61
% within group	8.2%	14.8%	77.0%	
Adjusted std. residual	(- 1.8)	(- 1.0)	(+ 2.3)	
Registered user	57	59	116	232
% within group	24.6%	25.4%	50.0%	
Adjusted std. residual	(+ 5.0)	(+ 3.3)	(- 6.6)	
Formerly reg. non-user	2	5	20	27
% within group	7.4%	18.5%	74.1%	
Adjusted std. residual	(- 1.3)	(- .1)	(+ 1.1)	
Totals	70	84	276	430

Note. The statement for this question is correct.

Registered users ($n = 232$) had the highest rate of accurately identifying the statement as correct at 24.6%, but also simultaneously had the highest rate of inaccurately identifying the statement as incorrect at 25.4%. Exactly one-half of registered users indicated that they did not know whether the statement is correct or incorrect. This suggests that the ways that @replies become visible to followers may not be well understood by registered users. Around three-fourths of every other user-type group indicated that they did not know the veracity of the statement. A subsequent chi-squared test showed a statistically

significant relationship between user-type and response choice, $X^2(6, N = 430) = 47.802$, $p < .05$.

Lastly, respondents were given the incorrect statement: "If you are using Twitter and someone sends you an @reply, but you are not following the user, the reply will still appear on your Tweets timeline." In actuality, unless you are following the user, the reply will instead appear on the "mentions" tab, but not in the timeline. Table 32 provides a breakdown of the responses to this statement, stratified by user-type.

Table 32

Crosstabulation of User-type and IFQ: @Replies and Timelines (Q19)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	21	1	88	110
% within group	19.1%	0.9%	80.0%	
Adjusted std. residual	(- 4.5)	(-3.1)	(+ 6.1)	
Unregistered user	27	0	35	62
% within group	43.5%	0.0%	56.5%	
Adjusted std. residual	(+ 1.2)	(- 2.5)	(+ .2)	
Registered user	105	33	95	233
% within group	45.1%	14.2%	40.8%	
Adjusted std. residual	(+ 3.8)	(+ 5.2)	(- 6.5)	
Formerly reg. non-user	6	0	20	26
% within group	23.1%	0.0%	76.9%	
Adjusted std. residual	(- 1.5)	(- 1.5)	(+ 2.3)	
Totals	159	34	238	431

Note. The statement for this question is incorrect.

Among registered users ($n = 233$), 45.1% inaccurately identified this statement as correct, only 14.2% accurately identified this statement as incorrect, and 40.8% indicated that they did not know whether the statement was correct or incorrect. While it seems as though this particular facet of information flows on Twitter is poorly understood by registered users, in re-examining the wording of the statement, it is possible that many registered users interpreted the "timeline" to also include the "mentions" tab. As a result,

this finding should be considered as suspect. Among other user-type groups, more than half indicated that they did not know whether the statement was accurate or inaccurate, and only 1 out of 198 accurately identified the statement as incorrect. A subsequent chi-squared test showed a statistically significant relationship between user-type and response choice, $X^2(6, N = 431) = 64.393$, $p < .05$.

Knowledge of Defaults among Sample

Following the questions on algorithms, participants were brought to the fourth of nine sets of questions regarding information flows on Twitter, questions about the defaults on the site. Based on the information flow framework developed in Chapter 4 (see Figure 5), defaults are part of the techno-cultural formation that constitute part of the "channel" of information flow. For this set, respondents were again instructed that they would be given a correct or incorrect statement and would then be asked to indicate that "Yes, this [statement] is correct," "No, this is incorrect," or "I don't know the answer to this question."

The first statement about defaults respondents were presented with is the incorrect statement: "By default, users can receive Direct Messages from any other users." This statement is incorrect because, by default, a user can only receive direct messages from other users that they follow. Table 33 provides a breakdown of the responses to this statement, stratified by user-type.

Table 33

Crosstabulation of User-type and IFQ: Direct Message Defaults (Q20)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	20	2	88	110
% within group	18.2%	1.8%	80.0%	
Adjusted std. residual	(- 4.0)	(- 4.0)	(+ 6.5)	
Unregistered user	11	3	48	62
% within group	17.7%	4.8%	77.4%	
Adjusted std. residual	(- 2.9)	(- 2.0)	(+ 4.1)	
Registered user	105	48	82	235
% within group	44.7%	20.4%	34.9%	
Adjusted std. residual	(+ 5.2)	(+ 5.3)	(- 8.4)	
Formerly reg. non-user	11	2	14	27
% within group	40.7%	7.4%	51.9%	
Adjusted std. residual	(+ .8)	(- .8)	(- .2)	
Totals	147	55	232	434

Note. The statement for this question is incorrect.

Registered users (n = 235) were the most successful at accurately identifying this statement as incorrect at a rate of 20.4%; however, registered users also had the highest rate of inaccurately identifying this statement as correct at a rate of 44.7%. Registered users indicated they did not know whether the statement is correct or incorrect at a rate of 34.9%. This suggests, as almost 80% of registered users could not accurately identify the statement as incorrect, that registered users may have poor understandings of who can send direct messages to whom by default. This weak rate of accurate identification was not only found among registered users though. More than half of every other user-type group indicated that they did not know whether this statement was correct or incorrect. A subsequent chi-squared test did, however, show a significant relationship between user-type and response to this question, $X^2(6, N = 434) = 82.525$, $p < .05$. These results suggest that the default information flows relating to direct messages may not be well understood

across all user-types, but that specifically, registered users may have a particularly high degree of incorrect knowledge and uncertainty.

Next, respondents were given the incorrect statement: "By default, your Tweets are "protected," meaning that, unless you change a setting, your Tweets will only be accessible to your followers." This statement is incorrect as by default, Tweets are public. Table 34 provides a breakdown of the responses to this statement, stratified by user-type.

Table 34

Crosstabulation of User-type and IFQ: Protected Account Defaults (Q21)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	6	17	87	110
% within group	5.5%	15.5%	79.1%	
Adjusted std. residual	(- 2.4)	(- 5.5)	(+ 6.9)	
Unregistered user	7	13	42	62
% within group	11.3%	21.0%	67.7%	
Adjusted std. residual	(- .2)	(- 2.9)	(+ 2.9)	
Registered user	36	120	78	234
% within group	15.4%	51.3%	33.3%	
Adjusted std. residual	(+ 2.3)	(+ 6.6)	(- 7.9)	
Formerly reg. non-user	3	11	13	27
% within group	11.1%	40.7%	48.1%	
Adjusted std. residual	(- .1)	(+ .4)	(- .3)	
Totals	52	161	220	433

Note. The statement for this question is incorrect.

Just over half of the registered users (n = 234) accurately identified this statement as incorrect (51.3%). One-third of registered users indicated they did not know whether this was a correct or incorrect statement, and 15.4% inaccurately identified this as a correct statement. The finding that just shy of one half of registered users could not accurately identify this statement as incorrect is particularly troubling, given the statement by Twitter founder Biz Stone in a blog post announcing the Twitter was giving the archive of all Tweets ever generated to the Library of Congress that, "… most of these tweets are

created with the intent that they will be publicly available" (Stone, 2010b, para. 2). Given that just under half of registered users sampled were either unsure or incorrect in their understanding of whether Tweets are "protected" by default, the assertion of intent may be problematic. Non-users, unregistered users, and formerly registered non-users also had high rates of indicating they did not know whether the statement is correct or incorrect (79.1%, 67.7% and 48.1% respectively). A subsequent chi-squared analysis found a statistically significant relationship between user-type and response, $X^2(6, N = 433) = 72.072$, $p < .05$.

Next, respondents were given the correct statement, "Unless you make changes to the default choices on your Twitter settings page, Twitter tailors its suggestions of the people you might enjoy following based on your recent visits to other websites that have integrated Twitter buttons or widgets." Table 35 provides a breakdown of the responses to this statement.

Table 35

Crosstabulation of User-type and IFQ: Default Tailoring People Suggested (Q22)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	26	1	82	109
% within group	23.9%	0.9%	75.2%	
Adjusted std. residual	(- 4.6)	(- 1.4)	(+ 5.0)	
Unregistered user	24	0	38	62
% within group	38.7%	0.0%	61.3%	
Adjusted std. residual	(- .7)	(- 1.4)	(+ 1.1)	
Registered user	118	10	106	234
% within group	50.4%	4.3%	45.3%	
Adjusted std. residual	(+ 3.6)	(+ 2.1)	(- 4.2)	
Formerly reg. non-user	16	1	10	27
% within group	59.3%	3.7%	37.0%	
Adjusted std. residual	(+ 1.8)	(+ .3)	(- 1.9)	
Totals	184	12	236	432

Note. The statement for this question is correct.

More than half of all registered users (n = 234) accurately identified this statement as correct, only 4.3% inaccurately identified this statement as incorrect, and 45.3% indicated that they did not know whether the statement is correct or incorrect. This suggests that while there is little outright misunderstanding of the fact that Twitter, by default, tailors its suggestions of people to follow based on website visits among registered users, there does appear to be a similar amount of accurate understanding and uncertainty about this particular setting. Interestingly, among formerly registered non-users (n = 27), there is a slightly higher percentage of individuals who accurately identified this statement as correct (59.3%) than among registered users. However, among unregistered users (n = 62) and non-users (n = 109), the rate of accurate identification dropped to 38.7% and 23.9% respectively, and majority of unregistered users and non-users indicated that they did not know whether this statement is correct or incorrect (75.2% and 61.3% respectively). A subsequent chi-squared analysis found a statistically significant relationship between user-type and response, $X^2(6, N = 432) = 33.617$, $p < .05$.

Next, respondents were given the incorrect statement: "Unless you make changes to the default choices on your Twitter settings page, location information (such as GPS coordinates) about where you are tweeting from is automatically made publicly accessible along with your Tweets." This statement is incorrect because users must consciously choose to change defaults in order to attach location information, such as GPS coordinates, along with Tweets. Table 36 provides a breakdown of the responses to this statement, stratified by user-type.

Table 36

Crosstabulation of User-type and IFQ: Defaults and Geo-location on Tweets (Q23)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	11	6	93	110
% within group	10.0%	5.5%	84.5%	
Adjusted std. residual	(- 2.7)	(- 3.8)	(+ 5.1)	
Unregistered user	9	9	44	62
% within group	14.5%	14.5%	71.0%	
Adjusted std. residual	(- .9)	(- .6)	(+ 1.2)	
Registered user	53	56	125	234
% within group	22.6%	23.9%	53.4%	
Adjusted std. residual	(+ 2.4)	(+ 4.1)	(- 5.2)	
Formerly reg. non-user	7	3	17	27
% within group	25.9%	11.1%	63.0%	
Adjusted std. residual	(+ 1.0)	(- .9)	(- .2)	
Totals	80	74	279	433

Note. The statement for this question is incorrect.

Registered users (n = 234) had the highest rate of accurately indicating that the statement is incorrect (23.9%); however, this group also had the highest rate of inaccurately identifying the statement as correct (22.6%). Over half of registered users indicated that they did not know whether this statement is correct or incorrect. This suggests that registered users may have a fair amount of uncertainty or outright misunderstandings of whether or not Twitter collects GPS information along with Tweets by default. More than 60% of other user-type groups indicated that they "did not know" whether this statement was correct or incorrect, which suggests the confusion about the default settings for geo-location information being uploaded to Twitter extends beyond registered users. A subsequent chi-squared analysis did, however, find a statistically significant relationship between user-type and response, $X^2(6, N = 432) = 35.480$, $p < .05$.

For the final prompt regarding default settings on Twitter, respondents were given the correct statement: "Unless you make changes to the default choices on your Twitter

settings page, Twitter tailors the advertisements you receive based on the information it gets about you from third-parties." Table 37 provides a breakdown of the responses to this statement, stratified by user-type.

Table 37

Crosstabulation of User-type and IFQ: Defaults and Tailored Advertisements (Q24)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	24	2	84	110
% within group	21.8%	1.8%	76.4%	
Adjusted std. residual	(- 3.9)	(- .4)	(+ 4.0)	
Unregistered user	22	1	39	62
% within group	35.5%	1.6%	62.9%	
Adjusted std. residual	(- .4)	(- .4)	(+ .5)	
Registered user	104	7	124	235
% within group	44.3%	3.0%	52.8%	
Adjusted std. residual	(+ 3.1)	(+ 1.0)	(- 3.4)	
Formerly reg. non-user	13	0	14	27
% within group	48.1%	0.0%	51.9%	
Adjusted std. residual	(+ 1.2)	(- .8)	(- .9)	
Totals	163	10	261	434

Note. The statement for this question is correct.

More than half of each user-type group indicated that they did not know the answer to this question. However, few respondents (10 of 434) inaccurately identified the statement as incorrect. Formerly registered non-users (n = 27) and registered users (n = 235) had the highest rate of accurately identifying the statement as correct (48.1% and 44.3% respectively). This suggests that there is not great outright misunderstanding about the defaults around tailored advertisements, but that this is more widespread uncertainty than accurate knowledge among the sample. As a result of violations to the assumptions that undergird the use of chi-squared tests for this question (as three cells (25.0%) had expected count less than five) a chi-squared test could not be used to test whether a significant relationship exists between the two categories.

Knowledge of Interfaces among Sample

Following the questions on defaults, participants were brought to the fifth of nine sets of questions regarding information flows on Twitter; questions about the interfaces on the site. Based on the information flow framework developed in Chapter 4 (see Figure 5), data and metadata structures are part of the techno-cultural formation that constitute part of the "channel" of information flow. For this set, respondents were again instructed that they would be given a correct or incorrect statement and would then be asked to indicate that "Yes, this [statement] is correct," "No, this is incorrect," or "I don't know the answer to this question."

The first statement users were presented with is the correct statement: "Twitter offers interfaces for programmers and application developers called the Twitter APIs, which can be used to make applications, websites, widgets, and other projects that interact with Twitter." Table 38 provides a breakdown of the responses to this statement, stratified by user-type.

Table 38

Crosstabulation of User-type and IFQ: Twitter Offers APIs (Q25)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	14	1	95	110
% within group	12.7%	0.9%	86.4%	
Adjusted std. residual	(-3.9)	(- .3)	(+ 3.9)	
Unregistered user	22	1	39	62
% within group	35.5%	1.6%	62.9%	
Adjusted std. residual	(+ 1.7)	(+ .4)	(- 1.7)	
Registered user	73	3	157	233
% within group	31.3%	1.3%	67.4%	
Adjusted std. residual	(+ 2.3)	(+ .3)	(- 2.3)	
Formerly reg. non-user	7	0	20	27
% within group	25.9%	0.0%	74.1%	
Adjusted std. residual	(- .1)	(- .6)	(+ .2)	
Totals	116	5	311	432

Note. The statement for this question is correct

Across all user-type groups, 311 of 432 respondents (71.9%) indicated that that they did not know if this statement was correct or incorrect. Registered users (n = 233) tended slightly below this, at a rate of 67.4% within the group. Unregistered users (n = 62) and registered users had the highest rate of accurately identifying this statement as correct, at 35.5% and 31.3% respectively. There were only 5 of 432 respondents who inaccurately identified this statement as incorrect. What this suggests is that there is not extensive misunderstanding of whether or not there are separate interfaces that programmers have access to for applications, but rather a general uncertainty or lack of knowledge among all user-groups. This uncertainty is highest among non-users (n = 110), as 86.4% indicated they did not know if the statement was correct or incorrect. As a result of violations to the assumptions that undergird the use of chi-squared tests for this question (as four cells (33.3%) had expected count less than five) a chi-squared test could not be used to test whether a significant relationship exists between the two categories.

Next, respondents were given the incorrect statement: "Old Tweets are automatically deleted from Twitter's servers after 2 years." This is an incorrect statement because Twitter does not automatically delete old Tweets after any period of time. Table 39 provides a breakdown of the responses to this statement, stratified by user-type.

Table 39

Crosstabulation of User-type and IFQ: Does Twitter Delete Old Tweets (Q26)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	3	7	99	109
% within group	2.8%	6.4%	90.8%	
Adjusted std. residual	(- .5)	(- 5.3)	(+ 5.3)	
Unregistered user	4	11	47	62
% within group	6.5%	17.7%	75.8%	
Adjusted std. residual	(+ 1.4)	(- 1.5)	(+ .9)	
Registered user	8	81	145	234
% within group	3.4%	34.6%	62.0%	
Adjusted std. residual	(- .1)	(+ 4.7)	(- 4.5)	
Formerly reg. non-user	0	11	16	27
% within group	0.0%	40.7%	59.3%	
Adjusted std. residual	(- 1.0)	(+ 1.9)	(- 1.4)	
Totals	15	110	307	432

Note. The statement for this question is incorrect.

Of registered users (*n* = 234), 34.6% were able to accurately identify this statement as incorrect. Only 3.4% inaccurately identified this statement as correct; however, 62.0% indicated they did not know whether this statement was correct or incorrect. This suggests that while there is not a high degree of outright inaccurate knowledge about Twitter's Tweet retention practices among registered users, there does appear to be a large degree of uncertainty about what those retention practices are. Interestingly, formerly registered users (*n* = 27) had the highest rates of accurately indicating this statement is incorrect at 40.7%. Among non-users (*n* = 109) and unregistered users (*n* = 62), over three-quarters indicated that they did not know whether the statement is correct or incorrect. A chi-squared test revealed a statistically significant relationship between user-status and response, $X^2(6, N = 432) = 39.273$, $p < .05$.

Respondents were next given the incorrect statement: "When you visit a website with Twitter buttons or widgets like the "Tweet This" button, Twitter does not receive information about that visit unless you click on the button or widget." This statement is

incorrect because Twitter receives information about the visit, regardless of whether or not a user clicks on a button or widget. Table 40 provides a breakdown of the responses to this statement, stratified by user-type.

Table 40

Crosstabulation of User-type and IFQ: Twitter Receives Data about Site Visits from Buttons and Widgets (Q27)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	12	14	83	109
% within group	11.0%	12.8%	76.1%	
Adjusted std. residual	(- 1.4)	(- .8)	(+ 1.7)	
Unregistered user	13	10	39	62
% within group	21.0%	16.1%	62.9%	
Adjusted std. residual	(+ 1.4)	(+ .2)	(- 1.2)	
Registered user	39	36	158	233
% within group	16.7%	15.5%	67.8%	
Adjusted std. residual	(+ 1.0)	(+ .1)	(-.8)	
Formerly reg. non-user	1	6	19	26
% within group	3.8%	23.1%	73.1%	
Adjusted std. residual	(- 1.7)	(+ 1.1)	(+ .4)	
Totals	65	66	299	430

Note. The statement for this question is incorrect.

Between 62 and 76 percent of respondents in every user-type group indicated that they did not know whether this statement is correct or incorrect. Across all respondents, almost an equal number inaccurately indicated that the statement is correct (65 respondents of 430, or 15.1%) as accurately identified the statement as incorrect (66 of 430, or 15.3%). In fact, a chi-squared test revealed no statistically significant relationship between user-status and response, $X^2(6, N = 430) = 7.918$, p = .244.

Next, respondents were given the correct statement: "Twitter offers a search interface to programmers that allows them to search for Tweets by GPS data (latitude, longitude and radius area), and will attempt to find Tweets created by users whose profile location matches the search parameters." Table 41 provides a breakdown of the responses to this statement, stratified by user-type.

Table 41

Crosstabulation of User-type and IFQ: Location-based Search API (Q28)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	10	1	97	108
% within group	9.3%	0.9%	89.8%	
Adjusted std. residual	(- 3.3)	(- 1.1)	(+ 3.6)	
Unregistered user	11	3	48	62
% within group	17.7%	4.8%	77.4%	
Adjusted std. residual	(- .6)	(+ 1.4)	(+ .1)	
Registered user	60	6	166	232
% within group	25.9%	2.6%	71.6%	
Adjusted std. residual	(+ 3.0)	(+ .4)	(- 3.0)	
Formerly reg. non-user	7	0	20	27
% within group	25.9%	0.0%	74.1%	
Adjusted std. residual	(+ .7)	(- .8)	(- .4)	
Totals	88	10	331	429

Note. The statement for this question is correct.

Registered users (n = 232) had the highest rates of accurately identifying this statement as correct among all user-type groups, at a rate of 25.9%. While only 6 of 232 registered users (2.6%) inaccurately identified this statement as incorrect, a clear majority of registered users (71.6%) did not know whether this statement is correct or incorrect. Given that this particular search function is not part of the standard web-based interface for non-API users, it is perhaps unsurprising that registered users would be unaware of this particular facet of information flows on the platform. Uncertainty about the veracity of this statement was also widespread among other user-type groups: at least 74% of all other groups indicated that they did not know. That being said, few respondents inaccurately identified the statement as incorrect (only 10 of 429 including registered users did so). As a result of violations to the assumptions that undergird the use of chi-squared tests for this question (as three cells (25.0%) had expected count less than five) a chi-squared test could not be used to test whether a significant relationship exists between the two categories.

Respondents were next presented with the incorrect statement, "There is no way for a user to retrieve all of the Tweets they have ever created." This statement is incorrect as Twitter introduced a tool in late 2012 that allows a user to bulk retrieve all the messages that they have created (Lynley, 2012). Table 42 provides a breakdown of the responses to this statement, stratified by user-type.

Table 42

Crosstabulation of User-type and IFQ: Retrieving Old Tweets (Q29)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	5	10	95	110
% within group	4.5%	9.1%	86.4%	
Adjusted std. residual	(+ .1)	(- 6.1)	(+ 5.9)	
Unregistered user	3	17	42	62
% within group	4.8%	27.4%	67.7%	
Adjusted std. residual	(+ .2)	(- 1.0)	(+ .9)	
Registered user	10	104	119	233
% within group	4.3%	44.6%	51.1%	
Adjusted std. residual	(- .1)	(+ 5.6)	(- 5.4)	
Formerly reg. non-user	1	11	15	27
% within group	3.7%	40.7%	55.6%	
Adjusted std. residual	(- .2)	(+ .9)	(- .8)	
Totals	19	142	271	432

Note. The statement for this question is incorrect.

Almost half (44.6%) of registered users (n = 233) accurately identified this statement as incorrect. Less than 5% of registered users inaccurately identified this statement as correct, though over 51% indicated they did not know whether this was a correct or incorrect statement. This suggests that while there is not much inaccurate knowledge of the fact that Twitter offers a tool for the bulk retrieval of old Tweets, that there may be as much uncertainty among registered users as accurate knowledge. Among other user-type groups, again, there is not a high degree of inaccurate identification of the statement as correct, but the rates of indicating uncertainty about whether the statement is correct or incorrect are higher. As a result of violations to the assumptions that undergird the use of

chi-squared tests for this question (as three cells (25.0%) had expected count less than five) a chi-squared test could not be used to test whether a significant relationship exists between the two categories. Given that this feature of Twitter was only implemented in late 2012, this would be a question worth returning to as part of a more longitudinal analysis.

The last statement respondents were presented within this section is the correct statement, "When you visit a website with Twitter buttons or widgets like the "Tweet This" button, Twitter may receive information about the URL (web-address) of the website you are visiting." Table 43 provides a breakdown of the responses to this statement, stratified by user-type.

Table 43

Crosstabulation of User-type and IFQ: Buttons and Widgets Leading to URL (Q30)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	31	3	76	110
% within group	28.2%	2.7%	69.1%	
Adjusted std. residual	(- 3.7)	(+ .0)	(+ 3.7)	
Unregistered user	24	1	37	62
% within group	38.7%	1.6%	59.7%	
Adjusted std. residual	(- .8)	(- .6)	(+ 1.0)	
Registered user	119	8	106	233
% within group	51.1%	3.4%	45.5%	
Adjusted std. residual	(+ 3.5)	(+ .9)	(- 3.8)	
Formerly reg. non-user	13	0	14	27
% within group	48.1%	0.0%	51.9%	
Adjusted std. residual	(+ .5)	(- .9)	(- .2)	
Totals	187	12	233	432

Note. The statement for this question is correct.

Slightly over half of registered users (n = 233) were able to accurately identify this statement as correct (51.1%), while 3.4% inaccurately identified this statement as incorrect, and 45.5% indicated that they did not know whether the statement is correct or incorrect. This suggests that while there is not much inaccurate knowledge of the fact that

Twitter receives the URLs of third-party websites on sites with buttons and widgets, there may be as much uncertainty as accurate knowledge. Overall, few respondents (only 12 of 432 in total) inaccurately identified the statement as incorrect, whereas a majority (233 of 432) instead indicated they did not know whether the statement was true or false. As a result of violations to the assumptions that undergird the use of chi-squared tests for this question (as three cells (25.0%) had expected count less than five) a chi-squared test could not be used to test whether a significant relationship exists between the two categories.

Knowledge of Business Practices among Sample

Following the questions on interfaces, participants were brought to the sixth of nine sets of questions regarding information flows on Twitter, questions about the business practices of Twitter. Based on the information flow framework developed in Chapter 4 (see Figure 5), business practices are part of the socio-economic formation that constitute and shape information flows on Twitter. A different question and response structure was used for this set. Respondents were told: "Please identify which of the following are ways that Twitter generates revenue (you can make multiple selections for this question)." Respondents were then given nine selection options, as follows:

1. Twitter generates revenue by selling advertising in the form of "promoted Tweets" that appear in users' timelines. [correct]
2. Twitter generates revenue by charging people for verified accounts. [incorrect]
3. Twitter generates revenue by selling access to the full stream of real-time Tweets created by users to third-parties. [correct]
4. Twitter generates revenue by charging other websites to put Twitter buttons and widgets on their sites. [incorrect]
5. Twitter generates revenue by selling advertising in the form of "promoted trends" that appear in the "Trending Now" section of the site. [correct]

6. Twitter generates revenue by charging news outlets like CNN every time they talk about things that have happened on Twitter. [incorrect]
7. Twitter generates revenue by selling advertising in the form of "promoted accounts" that are recommended to users. [correct]
8. Twitter generates revenue by charging websites that have Twitter buttons every time a visitor to those websites clicks on a "Tweet This" button. [incorrect]
9. I don't know the answer to this question.

Table 44 provides a breakdown of the responses to the correct first option, "Twitter generates revenue by selling advertising in the form of "promoted Tweets" that appear in users' timelines," stratified by user-type.

Table 44

Crosstabulation of User-type and IFQ: Revenue Generation, Promoted Tweets (Q31-O1)

Status	Unselected	Selected	Totals
Non-user	67	43	110
% within group	60.9%	39.1%	
Adjusted std. residual	(+ 5.9)	(- 5.9)	
Unregistered user	23	39	62
% within group	37.1%	62.9%	
Adjusted std. residual	(+ 0)	(+ 0)	
Registered user	63	172	235
% within group	26.8%	73.2%	
Adjusted std. residual	(-4.9)	(+ 4.9)	
Formerly reg. non-user	9	18	27
% within group	33.3%	66.7%	
Adjusted std. residual	(- .4)	(+ .4)	
Totals	162	272	434

Note. The statement for this question is correct and thus should be selected.

Almost three-quarters of registered users (n = 235) accurately indicated that Twitter generates revenue by selling "promoted Tweets" (73.2%). More than half (62.9%) of unregistered users (n = 62) and formerly registered non-users (n = 27) (66.7%) also accurately indicated that Twitter generates revenue by selling "promoted Tweets." Only 39.1% of non-users (n = 110) accurately identified the sale of promoted Tweets as a way that Twitter generates revenue. A subsequent chi-squared test did find a statistically significant relationship between user-status and response to this prompt, $X^2(3, N = 434)$ =

37.448, p < .05. As Twitter clearly indicates when a Tweet is a "promoted Tweet" in a user's timeline, it does make sense that respondents who may have been directly exposed to these types of Tweets previously would more readily identify them as a way Twitter generates revenue.

Table 45 provides a breakdown of the responses to the incorrect second option, "Twitter generates revenue by charging people for verified accounts," stratified by user-type.

Table 45

Crosstabulation of User-type and IFQ: Revenue Generation, Verified Accounts (Q31-O2)

Status	Unselected	Selected	Totals
Non-user	106	4	110
% within group	96.4%	3.6%	
Adjusted std. residual	(+2.0)	(- 2.0)	
Unregistered user	59	3	62
% within group	95.2%	4.8%	
Adjusted std. residual	(+ 1.0)	(- 1.0)	
Registered user	210	25	235
% within group	89.4%	10.6%	
Adjusted std. residual	(- 2.1)	(+ 2.1)	
Formerly reg. non-user	24	3	27
% within group	88.9%	11.1%	
Adjusted std. residual	(- .6)	(+ .6)	
Totals	399	35	434

Note. The statement for this question is incorrect and thus should be unselected.

Of the 434 total respondents, only 35 inaccurately identified this statement as a way that Twitter generates revenue (8.1%). The highest rate of inaccurate identification occurred within the formerly registered non-user group and the registered user group (11.1% and 10.6% respectively). Unregistered users and non-users fared better at a 4.8% and 3.6% rate of inaccurate identification. However, a chi-squared test revealed no statistically significant relationship between user-type and response to this question, $X^2(3, N = 434) = 37.448$, p = .102.

Table 46 provides a breakdown of the responses to the correct third option, "Twitter generates revenue by selling access to the full stream of real-time Tweets created by users to third-parties," stratified by user-type.

Table 46

Crosstabulation of User-type and IFQ: Revenue Generation, Firehose Access (Q31-O3)

Status	Unselected	Selected	Totals
Non-user	99	11	110
% within group	90.0%	10.0%	
Adjusted std. residual	(+ 2.7)	(- 2.7)	
Unregistered user	49	13	62
% within group	79.0%	21.0%	
Adjusted std. residual	(- .5)	(+ .5)	
Registered user	184	51	235
% within group	78.3%	21.7%	
Adjusted std. residual	(- 1.8)	(+ 1.8)	
Formerly reg. non-user	21	6	27
% within group	77.8%	22.2%	
Adjusted std. residual	(- .5)	(+ .5)	
Totals	353	81	434

Note. The statement for this question is correct and thus should be selected.

Just under one-quarter of registered users accurately indicated that Twitter generates revenue by selling access to the full stream of real-time Tweets created by users to third-parties (21.7%). Across the entire sample, only 81 of 434 respondents (18.7%) accurately identified this statement as correct. Identification of this statement as accurate by non-users was only 10.0%. However, a chi-squared test revealed no statistically significant relationship between user-type and accurate identification, $X^2(3, N = 434) = 7.310$, p = .063. Extrapolating from this, the fact that Twitter sells access to Tweets via its "firehose" may be poorly understood generally speaking, however, registered Twitter users show no statistically significant difference in their response patterns than the other user-type groups. This is perhaps unsurprising, given that other revenue generation methods, such as promoted Tweets, are far more visually accessible to users through common interaction interfaces, such as the timeline, whereas the "firehose" is not. This

also raises the question of whether Twitter's communication about this particular revenue generation practice to users is different from its communication about other revenue generation methods. More broadly, this finding also suggests that registered users may perceive Twitter as generating revenue through advertising, but may not widely recognize Twitter as generating revenue through the sale of access to user-generated content.

Table 47 provides a breakdown of the responses to the incorrect fourth option, "Twitter generates revenue by charging other websites to put Twitter buttons and widgets on their sites," stratified by user-type.

Table 47

Crosstabulation of User-type and IFQ: Revenue Generation, Buttons and Widgets (Q31-O4)

Status	Unselected	Selected	Totals
Non-user	88	22	110
% within group	80.0%	20.0%	
Adjusted std. residual	(+ 2.5)	(- 2.5)	
Unregistered user	41	21	62
% within group	66.1%	33.9%	
Adjusted std. residual	(- .8)	(+ .8)	
Registered user	160	75	235
% within group	68.1%	31.9%	
Adjusted std. residual	(- 1.2)	(+ 1.2)	
Formerly reg. non-user	17	10	27
% within group	63.0%	37.0%	
Adjusted std. residual	(- .9)	(+ .9)	
Totals	306	128	434

Note. The statement for this question is incorrect and thus should be unselected.

Of the 434 respondents, 128 inaccurately (29.5%) identified this statement as a way that Twitter generates revenue. Formerly registered non-users, unregistered users, and registered users had the highest rates of inaccurate identification at 37.0%, 33.9%, and 31.9% respectively. Non-users only misidentified this as a way Twitter generates revenue at a 20.0% rate. However, a subsequent chi-squared test revealed no statistically significant relationship between user-status and misidentification, $X^2(3, N = 434) = 6.740$,

p = .081. While almost one-third of respondents misidentified this as a way Twitter generates revenue, this misidentification was fairly uniform across multiple user-types.

Table 48 provides a breakdown of the responses to the correct fifth option, "Twitter generates revenue by selling advertising in the form of 'promoted trends' that appear in the "Trending Now" section of the site," stratified by user-type.

Table 48

Crosstabulation of User-type and IFQ: Revenue Generation, Promoted Trends (Q31-O5)

Status	Unselected	Selected	Totals
Non-user	72	38	110
% within group	65.5%	34.5%	
Adjusted std. residual	(+ 3.6)	(- 3.6)	
Unregistered user	32	30	62
% within group	51.6%	48.4%	
Adjusted std. residual	(+ .2)	(- .2)	
Registered user	103	132	235
% within group	43.8%	56.2%	
Adjusted std. residual	(- 3.0)	(+ 3.0)	
Formerly reg. non-user	12	15	27
% within group	44.4%	55.6%	
Adjusted std. residual	(- .6)	(+ .6)	
Totals	219	215	434

Note. The statement for this question is correct and thus should be selected.

For this response, more than half of registered users (n = 235) correctly identified "promoted trends" as a way that Twitter generates revenue (56.2%). Formerly registered users (n = 27) accurately identified this as a way Twitter generates revenue at a rate of 55.6%, unregistered users (n = 62) did so at a rate of 48.4%, and non-users (n = 110) did so at a rate of 34.5%. A subsequent chi-squared analysis revealed a statistically significant relationship between user-type and response, $X^2(3, N = 434) = 14.450$, $p < .05$. Similar to "promoted Tweets," Twitter clearly indicates when a trend is a "promoted trend" in the "Trending Topics" section of the interface. As a result, it does make sense that respondents who may have been directly exposed to these types of messages would

have a higher rate of accurate identification of this as a way Twitter generates revenue. Interestingly, among registered users, promoted Tweets were identified as a way Twitter generates revenue at a rate of 73.2%, whereas promoted trends dropped to 56.2%. This raises a question about why one type of promoted material may be better understood as a way that Twitter generates revenue than another, given that promoted trends, Tweets, and accounts were all introduced by Twitter roughly at the same time in 2010.

Table 49 provides a breakdown of the responses to the incorrect sixth option, "Twitter generates revenue by charging news outlets like CNN every time they talk about things that have happened on Twitter" stratified by user-type.

Table 49

Crosstabulation of User-type and IFQ: Revenue Generation, Media Charges (Q31-O6)

Status	Unselected	Selected	Totals
Non-user	103	7	110
% within group	93.6%	6.4%	
Adjusted std. residual	(+ 2.5)	(- 2.5)	
Unregistered user	52	10	62
% within group	83.9%	16.1%	
Adjusted std. residual	(- .7)	(+ .7)	
Registered user	198	37	235
% within group	84.3%	15.7%	
Adjusted std. residual	(- 1.6)	(+ 1.6)	
Formerly reg. non-user	23	4	27
% within group	85.2%	14.8%	
Adjusted std. residual	(- .2)	(+ .2)	
Totals	376	58	434

Note. The statement for this question is incorrect and thus should be unselected.

Overall, only 58 of 434 respondents inaccurately identified this statement as a way that Twitter generates revenue (13.4%). Unregistered users, registered users, and formerly registered non-users had the highest percentages of incorrectly identifying this statement as a way Twitter generates revenue, at 16.1%, 15.7% and 14.8% respectively. Non-users only inaccurately identified this as a way Twitter generates revenue at a rate of 6.4%.

However, a subsequent chi-squared test revealed no statistically significant relationship between user-type and response, $X^2(3, N = 434) = 6.265$, p = .099.

Table 50 provides a breakdown of the responses to the correct seventh option, "Twitter generates revenue by selling advertising in the form of "promoted accounts" that are recommended to users," stratified by user-type.

Table 50

Crosstabulation of User-type and IFQ: Revenue Generation, Promoted Accounts (Q31-O7)

Status	Unselected	Selected	Totals
Non-user	86	24	110
% within group	78.2%	21.8%	
Adjusted std. residual	(+ 4.4)	(- 4.4)	
Unregistered user	37	25	62
% within group	59.7%	40.3%	
Adjusted std. residual	(- .1)	(+ .1)	
Registered user	124	111	235
% within group	52.8%	47.2%	
Adjusted std. residual	(- 3.5)	(+ 3.5)	
Formerly reg. non-user	15	12	27
% within group	55.6%	44.4%	
Adjusted std. residual	(- .5)	(+ .5)	
Totals	262	172	434

Note. The statement for this question is correct and thus should be selected.

Overall, 172 of the 434 respondents correctly identified this as a way Twitter generates revenue (39.6%). Registered users, formerly registered non-users, and unregistered users had the highest rates of correct identification, at 47.2%, 44.4%, and 40.3% respectively. Non-users only correctly identified this as a way Twitter generates revenue at a rate of 21.8%. A subsequent chi-squared test revealed a statistically significant relationship between user-type and accurate identification, $X^2(3, N = 434) = 20.540$, $p < .05$. This finding does make sense as, similar to "promoted Tweets" and "promoted trends," Twitter clearly indicates when a recommended account is a promoted account in the timeline interface, and thus those who have been directly exposed to these types of promotions could more accurately identify them as a way Twitter generates revenue.

However, among registered users, promoted Tweets were identified as a way Twitter generates revenue at a rate of 73.2% and promoted trends at a rate of 56.2%, whereas promoted accounts fell to a level of just under half, 47.2%. This suggests that, even when visible as part of the user interface, some of Twitter's advertising-based revenue generation methods may be better understood by registered users than others.

Table 51 provides a breakdown of the responses to the incorrect eighth option, "Twitter generates revenue by charging websites that have Twitter buttons every time a visitor to those websites clicks on a "Tweet This" button," stratified by user-type.

Table 51

Crosstabulation of User-type and IFQ: Revenue Generation, "Tweet This" Clicks (Q31-O8)

Status	Unselected	Selected	Totals
Non-user	94	16	110
% within group	85.5%	14.5%	
Adjusted std. residual	(+ .8)	(- .8)	
Unregistered user	49	13	62
% within group	79.0%	21.0%	
Adjusted std. residual	(- .9)	(+ .9)	
Registered user	193	42	235
% within group	82.1%	17.9%	
Adjusted std. residual	(- .5)	(+ .5)	
Formerly reg. non-user	24	3	27
% within group	88.9%	11.1%	
Adjusted std. residual	(+ .8)	(- .8)	
Totals	360	74	434

Note. The statement for this question is incorrect and thus should be unselected.

Overall, 74 of 434 respondents inaccurately identified this as a way that Twitter generates revenue (17.1%). Unregistered users, registered users, and non-users had the highest rates of inaccurate identification, at 21.0%, 17.9%, and 14.5% respectively. However, a chi-squared test revealed no statistically significant relationship between user-status and inaccurate identification, $X^2(3, N = 434) = 1.946$, $p = .584$.

Table 52 provides a breakdown of the responses to the ninth and final option, "I don't know the answer to this question," stratified by user-type.

Table 52

Crosstabulation of User-type and IFQ: Revenue Generation, Do Not Know (Q31-O9)

Status	Unselected	Selected	Totals
Non-user	40	70	110
% within group	36.4%	63.6%	
Adjusted std. residual	(-5.7)	(+ 5.7)	
Unregistered user	39	23	62
% within group	62.9%	37.1%	
Adjusted std. residual	(+.6)	(- .6)	
Registered user	162	73	235
% within group	68.9%	31.1%	
Adjusted std. residual	(+ 4.4)	(- 4.4)	
Formerly reg. non-user	17	10	27
% within group	63.0%	37.0%	
Adjusted std. residual	(+ .4)	(- .4)	
Totals	258	176	434

Across all user-types, 176 of 434 respondents indicated that they did not know the answer to how Twitter generates revenue (40.6%). It should be noted however, that it was possible for respondents to select methods of revenue generation they believed to be correct and to select "do not know" due to non-exclusivity. As a result, do not know responses may have been selected alongside other responses. Non-users had the highest rate of indication, at 63.6%, and registered users had the lowest rate at 31.1%. A chi-squared test revealed a statistically significant relationship between user-type and selecting "I don't know the answer to this question," $X^2(3, N = 434) = 33.536$, $p < .05$. Despite this relationship, it is still somewhat surprising that almost one-third of registered users indicated that they did not know the answer to this question. Also, as more of a meta-commentary on the structure of the survey, in retrospect, it may have been more beneficial to structure the questions and answers to match the result of the survey, thus providing "do not know" selection for each option. Such a change would have allowed for greater granularity in the reporting of accurate/inaccurate/uncertain knowledge among user-types.

In addition to cross-tabulations of user-type and participant's responses regarding how Twitter generates revenue, it is useful to narrow in on each user-type and to highlight the most frequently occurring responses, sorted by rate of selection. Table 53 provides a list of rank ordered selections within each user group.

Table 53

Rank-order of Rate of Indication of how Twitter Generates Revenue within User-type Groups

	Non-Users	Unregistered Users	Registered Users	Formerly Registered Non-Users
1st	I don't know the answer to this question. (63.6%)	...by selling advertising in the form of "promoted Tweets" that appear in users' timelines. [correct] (62.9%)	...by selling advertising in the form of "promoted Tweets" that appear in users' timelines. [correct] (73.2%)	...by selling advertising in the form of "promoted Tweets" that appear in users' timelines. [correct] (66.7%)
2nd	...by selling advertising in the form of "promoted Tweets" that appear in users' timelines. [correct] (39.1%)	... by selling advertising in the form of "promoted trends" that appear in the "Trending Now" section of the site. [correct] (48.4%)	... by selling advertising in the form of "promoted trends" that appear in the "Trending Now" section of the site. [correct] (56.2%)	... by selling advertising in the form of "promoted trends" that appear in the "Trending Now" section of the site. [correct] (55.6%)
3rd	... by selling advertising in the form of "promoted trends" that appear in the "Trending Now" section of the site. [correct] (34.5%)	...by selling advertising in the form of "promoted accounts" that are recommended to users. [correct] (40.3%)	...by selling advertising in the form of "promoted accounts" that are recommended to users. [correct] (47.2%)	...by selling advertising in the form of "promoted accounts" that are recommended to users. [correct] (44.4%)
4th	...by selling advertising in the form of "promoted accounts" that are recommended to users. [correct] (21.8%)	I don't know the answer to this question. (37.1%)	...by charging other websites to put Twitter buttons and widgets on their sites. [incorrect] (31.9%)	...by charging other websites to put Twitter buttons and widgets on their sites. [incorrect] (37.0%)
5th	...by charging other websites to put Twitter buttons and widgets on their sites. [incorrect] (20.0%)	...by charging other websites to put Twitter buttons and widgets on their sites. [incorrect] (33.9%)	I don't know the answer to this question. (31.1%)	I don't know the answer to this question. (37.0%)
6th	...by charging websites that have Twitter buttons every time a visitor to those websites clicks on a "Tweet This" button. [incorrect] (14.5%)	...by selling access to the full stream of real-time Tweets created by users to third-parties. [correct] (21.0%)	...by selling access to the full stream of real-time Tweets created by users to third-parties. [correct] (21.7%)	...by selling access to the full stream of real-time Tweets created by users to third-parties. [correct] (22.2%)
7th	...by selling access to the full stream of real-time Tweets created by users to third-parties. [correct] (10.0%)	...by charging websites that have Twitter buttons every time a visitor to those websites clicks on a "Tweet This" button. [incorrect] (21.0%)	...by charging websites that have Twitter buttons every time a visitor to those websites clicks on a "Tweet This" button. [incorrect] (17.9%)	... by charging news outlets like CNN every time they talk about things that have happened on Twitter. [incorrect] (14.8%)
8th	... by charging news outlets like CNN every time they talk about things that have happened on Twitter. [incorrect] (6.4%)	... by charging news outlets like CNN every time they talk about things that have happened on Twitter. [incorrect] (16.1%)	... by charging news outlets like CNN every time they talk about things that have happened on Twitter. [incorrect] (15.7%)	...by charging people for verified accounts. [incorrect] (11.1%)
9th	...by charging people for verified accounts. [incorrect] (3.6%)	...by charging people for verified accounts. [incorrect] (4.8%)	...by charging people for verified accounts. [incorrect] (10.6%)	...by charging websites that have Twitter buttons every time a visitor to those websites clicks on a "Tweet This" button. [incorrect] (11.1%)

Note. Green cells indicate a way that Twitter generates revenue, red cells indicate an inaccurate way Twitter generates revenue, and blue cells indicate "don't know" responses.

As Table 53 illustrates, non-users most frequently selected "I don't know the answer to this question" at a rate of 63.6%. However, following this, non-users accurately indicated that promoted Tweets, promoted trends, and promoted accounts are ways that Twitter generates revenue as the second, third, and fourth most selected option (at rates of 39.1%, 34.5%, and 21.8%, respectively). However, the fifth and sixth most selected options are not real ways that Twitter generates revenue. The final correct response (sale of access to the firehose) was the seventh most selected option, with only 10.0% of non-user respondents having indicated that this is a way Twitter generated revenue. Overall, these results suggest that a majority of non-users have uncertainty about how Twitter generates revenue, though promoted Tweets and promoted trends are recognized as ways Twitter generates revenue by more than one-third of the sampled non-users. Promoted accounts and the sale of access to Tweets via the firehose appear to be more unfamiliar to non-users.

Unregistered users accurately identified that promoted Tweets, promoted trends, and promoted accounts as ways that Twitter generates revenue as the first, second, and third most selected option (at rates of 62.9%, 48.4%, and 40.3%, respectively). The fourth most selected option was the "I don't know the answer to this question" response, at a rate of 37.1%. However, the fifth most selected option is not a real way that Twitter generates revenue. The final correct response (sale of access to the firehose) was the sixth most selected option, with only 21.0% of respondents having indicated that this is a way Twitter generates revenue. Overall, these results suggest that a majority of sampled unregistered users could identify promoted Tweets, promoted trends, and promoted

accounts as ways that Twitter generates revenue; however, the sale of access to Tweets via the firehose to third-parties appears to be somewhat unfamiliar to unregistered users.

The rank ordering of indication rates among registered users closely follows that of unregistered users. Registered users accurately identified promoted Tweets, promoted trends, and promoted accounts as ways that Twitter generates revenue as the first, second, and third most selected option (at rates of 73.2%, 56.2%, and 47.2%, respectively). However, the fourth most selected option is not a real way that Twitter generates revenue. The fifth most selected option of registered users was the "I don't know the answer to this question" response at a rate of 31.1%. The final correct response (sale of access to the firehose) was the sixth most selected option, with only 21.7% of respondents indicating that this is a way Twitter generates revenue. Overall, these results suggest that a majority of sampled registered users could identify promoted Tweets, promoted trends, and promoted accounts as ways that Twitter generates revenue; however, again, the sale of access to Tweets to third-parties via the firehose appears to be less familiar to registered users.

Lastly, the rank ordering of indication rates among formerly registered non-users follows that of registered users. Formerly registered non-users accurately indicated that promoted Tweets, promoted trends, and promoted accounts are ways that Twitter generates revenue as the first, second, and third most selected option (at rates of 66.7%, 55.6%, and 44.4%, respectively). However, the fourth (tie with fifth) most selected option is not a real way that Twitter generates revenue and the fifth (tie with fourth) most selected option from unregistered users was the "I don't know the answer to this question" response at a rate of 37.0%. The final correct response (sale of access to

firehose) was the sixth most selected option, with only 22.2% of respondents indicating that this is a way Twitter generates revenue. Overall, these results suggest that a majority of formerly registered non-users could identify promoted Tweets, promoted trends, and promoted accounts as a way Twitter generates revenue; however, the sale of access to Tweets via the firehose appears to be familiar to only one in five formerly registered non-user respondents.

Knowledge of Governance among Sample

Following the questions on the business practices of Twitter, participants were brought to the seventh of nine sets of questions regarding information flows on Twitter, questions about the governance of the site. Based on the information flow framework developed in Chapter 4 (see Figure 5), governance practices are part of the socio-economic formation that constitute and shape information flows on Twitter. Respondents were told:

> For the following questions, we will be asking you a series of questions about Twitter. Please select what you believe to be the correct answer. Please do not use any external sources to help you answer these questions. You may choose "I don't know the answer to this question" if you feel that you do not know the answer.

Participants were then given a statement about protocols on Twitter that was either correct or incorrect, and were asked to indicate that "Yes, this [statement] is correct," "No, this is incorrect," or "I don't know the answer to this question."

The first statement presented to respondents is the correct statement: "Twitter has three documents that govern users on the site, the 'Terms of Service,' the 'Privacy Policy' and 'The Twitter Rules.'" Table 54 provides a breakdown of the responses to this statement.

Table 54

Crosstabulation of User-type and IFQ: Twitter's Three Governance Documents (Q32)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	18	2	90	110
% within group	16.4%	1.8%	81.8%	
Adjusted std. residual	(- 3.3)	(- 2.6)	(+ 4.5)	
Unregistered user	16	2	44	62
% within group	25.8%	3.2%	71.0%	
Adjusted std. residual	(- .5)	(- 1.4)	(+ 1.3)	
Registered user	84	25	126	235
% within group	35.7%	10.6%	53.6%	
Adjusted std. residual	(+ 3.6)	(+ 2.6)	(- 4.8)	
Formerly reg. non-user	6	4	17	27
% within group	22.2%	14.8%	63.0%	
Adjusted std. residual	(- .8)	(+ 1.5)	(- .1)	
Totals	124	33	277	434

Note. The statement for this question is correct.

Across all respondents (n = 434), 63.8% indicated that they "did not know" whether this was a correct or incorrect statement. While non-users (n = 110) had the highest rate of indicating they "don't know" (81.8%), more than half of all registered users (n = 235) also indicated that they did not know whether this statement was correct or incorrect. This is somewhat surprising as registered users, in theory, would have encountered these documents when they signed up for the service, as part of the registration process; however, as the findings from the media consumption and policy document reading section indicate, few registered users in the sample have ever read the documents beyond a cursory glance. Registered users did have the highest rates of accurately identifying the statement as correct at 35.7%, but also had the second highest rate of inaccurately indicating the statement is incorrect at 10.6%. A chi-squared analysis revealed a statistically significant relationship between user-status and response to this prompt, $X^2(6, N = 434) = 31.530$, $p < .05$.

The second statement users were presented with is the incorrect statement: "Twitter allows spam, abuse, phishing, and malware on its platform as long as it is marked as 'potentially sensitive.'" This statement is incorrect as Twitter does not allow these types of content, even if they are marked as potentially sensitive. Table 55 provides a breakdown of the responses to this statement, stratified by user-type.

Table 55

Crosstabulation of User-type and IFQ: Governance of Spam Content (Q33)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	9	17	84	110
% within group	8.2%	15.5%	76.4%	
Adjusted std. residual	(- .1)	(- 3.3)	(+ 3.2)	
Unregistered user	8	14	40	62
% within group	12.9%	22.6%	64.5%	
Adjusted std. residual	(+ 1.4)	(- 1.0)	(+ .1)	
Registered user	17	83	133	233
% within group	7.3%	35.6%	57.1%	
Adjusted std. residual	(- .8)	(+ 3.9)	(- 3.2)	
Formerly reg. non-user	2	6	19	27
% within group	7.4%	22.2%	70.4%	
Adjusted std. residual	(- .2)	(- .7)	(+ .7)	
Totals	36	120	276	432

Note. The statement for this question is incorrect.

Of all respondents (n = 432), 63.8% indicated that they did not know the answer to this question. While non-users (n = 110) had the highest rate of indicating they "don't know" (76.4%), more than half of registered users (n = 233) also indicated that they did not know whether this statement was correct or incorrect. It is perhaps not surprising that respondents would not know the answer to this question, given that Twitter actively deletes egregious violations of its content restriction rules (Stone, 2008) and users may not have encountered such content "in the wild." Registered users did, however, have the highest rates of accurately identifying the statement as correct at 35.6%. A chi-squared analysis revealed a statistically significant relationship between user-status and response to this prompt, $X^2(6, N = 432) = 18.514$, $p < .05$.

The third statement users were presented with is the correct statement: "Users of Twitter's APIs are forbidden from selling, renting, leasing, or redistributing access to the Twitter API or Twitter Content to any third party without prior written approval from Twitter." Table 56 provides a breakdown of the responses to this statement, stratified by user-type.

Table 56

Crosstabulation of User-type and IFQ: Governance of Twitter API Content Rebroadcasting (Q34)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	14	2	92	108
% within group	13.0%	1.9%	85.2%	
Adjusted std. residual	(- 2.6)	(- .2)	(+ 2.6)	
Unregistered user	14	0	47	61
% within group	23.0%	0.0%	77.0%	
Adjusted std. residual	(+ .2)	(- 1.2)	(+ .2)	
Registered user	60	6	169	235
% within group	25.5%	2.6%	71.9%	
Adjusted std. residual	(+ 2.0)	(+ .7)	(- 2.2)	
Formerly reg. non-user	6	1	20	27
% within group	22.2%	3.7%	74.1%	
Adjusted std. residual	(+ .1)	(+ .6)	(- .3)	
Totals	94	9	328	431

Note. The statement for this question is correct.

More than three-quarters of all respondents (n = 431) indicated that they did not know whether the statement about content redistribution was correct or incorrect (76.1%), although only 9 of 431 respondents (2.0%) inaccurately identified this statement as incorrect. Registered users (n = 235) had the highest rate of having accurately identified this statement as correct (25.5%) and the lowest rate among all user-type groups of having indicated that they did not know (71.9%). However, a chi-squared analysis revealed no statistically significant relationship between user-status and response to this question, $X^2(6, N = 431) = 9.025$, $p = .172$. The rules that govern what users of the APIs

can and cannot do with Twitter Content may not be well understood broadly, regardless of user-type.

The fourth statement users were presented with as part of this set is the correct statement: "Twitter advertisers are prohibited from offering adult or sexual products or services, drugs or drug paraphernalia, endangered species products, hate content, tobacco, unauthorized ticket sales, or weapons." Table 57 provides a breakdown of the responses to this statement, stratified by user-type.

Table 57

Crosstabulation of User-type and IFQ: Governance of Tweet Content (Q36)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	34	2	74	110
% within group	30.9%	1.8%	67.3%	
Adjusted std. residual	(- 2.9)	(- 1.4)	(+ 3.4)	
Unregistered user	28	3	31	62
% within group	45.2%	4.8%	50.0%	
Adjusted std. residual	(+ .4)	(+ .3)	(- .5)	
Registered user	114	12	108	234
% within group	48.7%	5.1%	46.2%	
Adjusted std. residual	(+ 2.7)	(+ 1.1)	(- 3.1)	
Formerly reg. non-user	9	1	17	27
% within group	33.3%	3.7%	63.0%	
Adjusted std. residual	(-1.0)	(- .1)	(+ 1.1)	
Totals	185	18	230	433

Note. The statement for this question is correct.

While 53.1% of all respondents (n = 433) indicated that they did not know whether this was a true or false statement, among registered users (n = 234), the rate of indicating "don't know" was only 46.2%. A majority of registered users were able to accurately identify this statement as correct (48.7%), and only 5.1% of registered users inaccurately identified this statement as incorrect. As a result of violations to the assumptions that undergird the use of chi-squared tests for this question (as three cells (25.0%) had

expected count less than five) a chi-squared test could not be used to test whether a significant relationship exists between the two categories.

The fifth and final statement users were presented with as part of this set is the correct statement: "All Tweets created by users with public accounts are given to the Library of Congress for archiving by Twitter." Table 58 provides a breakdown of the responses to this statement, stratified by user-type.

Table 58

Crosstabulation of User-type and IFQ: Tweets Archived by LoC (Q37)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	7	15	88	110
% within group	6.4%	13.6%	80.0%	
Adjusted std. residual	(- 3.5)	(+ .6)	(+ 2.5)	
Unregistered user	12	8	42	62
% within group	19.4%	12.9%	67.7%	
Adjusted std. residual	(+ .5)	(+ .2)	(- .5)	
Registered user	50	27	157	234
% within group	21.4%	11.5%	67.1%	
Adjusted std. residual	(+ 2.4)	(- .3)	(- 1.8)	
Formerly reg. non-user	6	2	19	27
% within group	22.2%	7.4%	70.4%	
Adjusted std. residual	(+ .7)	(- .8)	(+ .0)	
Totals	75	52	306	433

Note. The statement for this question is correct.

Across all respondents (n = 433), 70.7% indicated that they did not know whether this was a correct or incorrect statement. Among registered users (n = 234), the rate of selection of "don't know" fell to 67.1%. However, only 21.4% of registered users were able to accurately identify this as a correct statement, while 11.5% inaccurately identified it as an inaccurate statement. Non-users (n = 110) fared the worst on this question, with 80.0% having indicated they "don't know," 13.6% having inaccurately identified the statement as incorrect, and only 6.4% having accurately identified the statement as correct. The low rate of accurate identification among the sample, but particularly among the registered users is somewhat troubling, as this suggests that many users may not

know what happens to Tweets in the long-term. This is particularly the case also considered the high rate of "don't know" responses to questions about whether or not Twitter deletes Tweets. As a result of violations to the assumptions that undergird the use of chi-squared tests for this question (as three cells (25.0%) had expected count less than five) a chi-squared test could not be used to test whether a significant relationship exists between the two categories.

Knowledge of Users among Sample

Following the questions on the governance on Twitter, participants were brought to the eighth of nine sets of questions regarding information flows on Twitter, questions relating to the users of the site. Based on the information flow framework developed in Chapter 4 (see Figure 5), users are part of the techno-cultural formation that constitute the senders and receivers portion of information flows on Twitter. For this set of questions, respondents were again instructed that they would be given a correct or incorrect statement and would then be asked to indicate that "Yes, this [statement] is correct," "No, this is incorrect," or "I don't know the answer to this question."

Respondents were first presented with the incorrect statement: "Unregistered visitors to Twitter.com can still view publicly created Tweets but cannot use the "search" feature of the website." This statement is incorrect as unregistered visitors can use the search feature of the website in addition to being able to view non-protected Tweets. Table 59 provides a breakdown of the responses to this statement, stratified by user-type.

Table 59

Crosstabulation of User-type and IFQ: Unregistered User Access to Search (Q38)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	26	6	78	110
% within group	23.6%	5.5%	70.9%	
Adjusted std. residual	(- 2.4)	(- 3.7)	(+ 5.1)	
Unregistered user	32	11	19	62
% within group	51.6%	17.7%	30.6%	
Adjusted std. residual	(+ 3.4)	(+ .2)	(- 3.3)	
Registered user	78	48	108	234
% within group	33.3%	20.5%	46.2%	
Adjusted std. residual	(+ .1)	(+ 2.2)	(- 1.8)	
Formerly reg. non-user	7	8	12	27
% within group	25.9%	29.6%	44.4%	
Adjusted std. residual	(- .8)	(+ 1.8)	(- .6)	
Totals	143	73	217	433

Note. The statement for this question is incorrect.

The response to this question is somewhat vexing. Over half of unregistered users (n = 62) inaccurately indicated that this is a correct statement (51.6%). Registered users (n = 234) had the second highest rate of inaccurate identification at 33.3%. Meanwhile, formerly registered non-users (n = 27) had the highest rate of accurately identifying the statement at incorrect at 29.6%. A subsequent chi-squared test revealed a statistically significant relationship between user-type and response to this question, $X^2(6, N = 434) = 37.907$, $p < .05$.

Next, respondents were presented with the incorrect statement: "'Verified accounts' on Twitter are those for which the user has paid Twitter." This statement is incorrect as verified accounts are actually just accounts for which Twitter has verified the authenticity of the person or brand running the account. There is no exchange of money involved (as of the time of the study). Table 60 provides a breakdown of the responses to this statement, stratified by user-type.

Table 60

Crosstabulation of User-type and IFQ: Are Verified Accounts Paid For (Q39)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	7	15	87	109
% within group	6.4%	13.8%	79.8%	
Adjusted std. residual	(- 1.8)	(- 4.4)	(+ 5.3)	
Unregistered user	5	19	38	62
% within group	8.1%	30.6%	61.3%	
Adjusted std. residual	(- .8)	(+ .0)	(+ .5)	
Registered user	34	91	110	235
% within group	14.5%	38.7%	46.8%	
Adjusted std. residual	(+ 2.4)	(+ 3.9)	(- 5.2)	
Formerly reg. non-user	2	8	17	27
% within group	7.4%	29.6%	63.0%	
Adjusted std. residual	(- .6)	(- .1)	(+ .5)	
Totals	48	133	252	433

Note. The statement for this question is incorrect.

Registered users (n = 235) had the highest rate of accurately identifying this statement as incorrect at a rate of 38.7%. The rate of accurate identification dropped to roughly 30% for unregistered users (n = 62) and formerly registered non-users (n = 27), and fell to 13.8% among non-users (n = 109). Registered users did, however, also have the highest rate of inaccurately identifying the statement as correct at a rate of 14.5%. Overall, the majority of respondents in each user group indicated that they did not know whether the statement was incorrect or incorrect, with non-users having the highest "don't know" response rate at 79.8% and registered users having the lowest at 46.8%. A chi-squared test revealed a statistically significant relationship between user-type and response to this question, $X^2(6, N = 433) = 34.719$, $p < .05$.

Respondents were next presented with the incorrect statement: "Twitter's 'Certified Products' partners are all prohibited from reselling historical Twitter data, such as old Tweets, to third-parties." This statement is incorrect as some of Twitter's Certified Products partners (such as GNIP and DataSift) are allowed to resell historical Twitter

data. Table 61 provides a breakdown of the responses to this statement, stratified by user-type.

Table 61

Crosstabulation of User-type and IFQ: Do Certified Products Resell Tweets (Q40)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	7	3	98	108
% within group	6.5%	2.8%	90.7%	
Adjusted std. residual	(- 1.7)	(- 2.2)	(+ 2.9)	
Unregistered user	14	5	43	62
% within group	22.6%	8.1%	69.4%	
Adjusted std. residual	(+ 3.2)	(+ .1)	(- 2.6)	
Registered user	24	22	186	232
% within group	10.3%	9.5%	80.2%	
Adjusted std. residual	(- .4)	(+ 1.5)	(- .7)	
Formerly reg. non-user	2	3	22	27
% within group	7.4%	11.1%	81.5%	
Adjusted std. residual	(- .6)	(+ .7)	(+ .0)	
Totals	47	33	349	429

Note. The statement for this question is incorrect.

Across all respondents (n = 429), there was at 81.4% rate of selecting the "don't know" answer for this question. This was highest among non-users (n = 108) at 90.7% and lowest among unregistered users (n = 62) at 69.4%. Formerly registered non-users (n = 27) and registered users (n = 232) had the highest rate of accurately indicating that the statement is incorrect at 11.1% and 9.5% respectively. Unregistered users and registered users had the highest rates of incorrectly indicating the statement is correct at 22.6% and 10.3% respectively. A chi-squared analysis showed no statistically significant relationship between user-type and response, $X^2(6, N = 429) = 17.097$, $p = .172$. Accurate, inaccurate, and uncertain responses regarding what Twitter's Certified Products partners are and are not allowed to do with Twitter data were fairly uniformly across all user-type groups.

Next, respondents were given the correct statement: "If you have a "protected" account on Twitter and another user wants to follow you, you must approve them before

they can do so." Table 62 provides a breakdown of the responses to this statement, stratified by user-type.

Table 62

Crosstabulation of User-type and IFQ: Protected Accounts and Following Approval (Q41)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	34	2	72	108
% within group	31.5%	1.9%	66.7%	
Adjusted std. residual	(- 6.8)	(+ .2)	(+ 6.8)	
Unregistered user	36	0	26	62
% within group	58.1%	0.0%	41.9%	
Adjusted std. residual	(- .2)	(- 1.1)	(+ .5)	
Registered user	169	5	60	234
% within group	72.2%	2.1%	25.6%	
Adjusted std. residual	(+ 6.0)	(+ .9)	(- 6.3)	
Formerly reg. non-user	16	0	11	27
% within group	59.3%	0.0%	40.7%	
Adjusted std. residual	(+ .0)	(- .7)	(+ .2)	
Totals	255	7	169	431

Note. The statement for this question is correct.

A majority of registered users (n = 234) accurately identified this statement as correct (72.2%), suggesting that what protected accounts afford in terms of following permissions is fairly well understood by registered users. Among other user-type groups, the rates of accurate identification of the statement as correct slipped slightly, but remained at 50% for unregistered users (n = 62) and formerly registered non-users (n = 27). The rate of accurate identification fell to 31.5% among non-users (n = 108), "Don't know" was the second most common response across the entire population, with 66.7% of non-users, 41.9% of unregistered users, 40.7% of formerly registered non-users, and 25.6% of registered users selecting this option within each respective group. Across the entire pool of respondents, only 7 of 431 (1.6%) inaccurately identified this statement as incorrect. As a result of violations to the assumptions that undergird the use of chi-squared tests for this question (as four cells (33.0%) had expected count less than five) a

chi-squared test could not be used to test whether a significant relationship exists between the two categories.

Respondents were next presented with the correct statement: "'Verified accounts' on Twitter are those for which the identities of the individuals or brands they represent have been authenticated." Table 63 provides a breakdown of the responses to this statement, stratified by user-type.

Table 63

Crosstabulation of User-type and IFQ: Verified Account Authentication (Q42)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	29	1	78	108
% within group	26.9%	0.9%	72.2%	
Adjusted std. residual	(- 6.9)	(- .8)	(+ 7.2)	
Unregistered user	31	2	29	62
% within group	50.0%	3.2%	46.8%	
Adjusted std. residual	(- .9)	(+ .9)	(+ .7)	
Registered user	163	5	63	231
% within group	70.6%	2.2%	27.3%	
Adjusted std. residual	(+ 6.8)	(- .5)	(- 7.0)	
Formerly reg. non-user	14	0	13	27
% within group	51.9%	0.0%	48.1%	
Adjusted std. residual	(- .4)	(- .7)	(+ .6)	
Totals	237	8	183	428

Note. The statement for this question is correct.

A majority of respondents within the registered users group (n = 231), formerly registered non-user (n = 27) and unregistered user groups (n = 62) accurately identified this statement as correct (70.6%, 51.9% and 50.0%, respectively). "Don't know" was the second most common response across the entire population (n = 428), with 72.2% of non-users, 48.1% of formerly registered non-users, 46.8% of unregistered users, and 25.6% of registered users having selected this option within each respective group. Across the entire pool of respondents, only 8 of 428 (1.9%) inaccurately identified this statement as incorrect. As a result of violations to the assumptions that undergird the use

of chi-squared tests for this question (as three cells (25.0%) had expected count less than five) a chi-squared test could not be used to test whether a significant relationship exists between the two categories.

Respondents were next given the correct statement: "Advertisers who purchase 'promoted Tweets' on Twitter receive information about the number of users that have been served that Tweet, the number of clicks on a promoted Tweet (and information on which piece of the Tweet users clicked on), the number of times a Tweet has been retweeted, and the number of times a promoted Tweet has been replied to." Table 64 provides a breakdown of the responses to this statement, stratified by user-type.

Table 64

Crosstabulation of User-type and IFQ: Info Advertisers Receive (Q43)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	29	1	79	109
% within group	26.6%	0.9%	72.5%	
Adjusted std. residual	(- 3.8)	(+ .3)	(+ 3.7)	
Unregistered user	29	1	32	62
% within group	46.8%	1.6%	51.6%	
Adjusted std. residual	(+ .8)	(+ .9)	(- 1.0)	
Registered user	115	1	117	233
% within group	49.4%	0.4%	50.2%	
Adjusted std. residual	(+ 3.4)	(- .7)	(- 3.2)	
Formerly reg. non-user	8	0	19	27
% within group	29.6%	0.0%	70.4%	
Adjusted std. residual	(- 1.3)	(- .4)	(+ 1.4)	
Totals	181	3	247	431

Note. The statement for this question is correct.

A majority of respondents in each user group that they did not know whether this statement is correct or incorrect. Non-users (n = 109) had the highest rate of selecting "don't know" at a rate of 72.5%, and registered users had the lowest at 50.2%. Registered users (n = 233) and unregistered users (n = 62) had the highest rates of accurately identifying this statement as correct at 49.4% and 46.8% respectively. Formerly registered non-users (n = 27) and non-users had the lowest rate of accurately selecting

"the statement is correct" at 29.6% and 26.6% respectively. Across the entire pool of respondents, only 3 of 431 inaccurately indicated the statement is incorrect (less than 1%). As a result of violations to the assumptions that undergird the use of chi-squared tests for this question (as four cells (33.3%) had expected count less than five) a chi-squared test could not be used to test whether a significant relationship exists between the two categories.

Respondents were next presented with the seventh and final statement of this section, the incorrect statement: "Information about the number of Tweets, number of photos, number of followers and followees, and number of favorites of users with protected accounts is not publicly accessible information." In fact, despite it coming from "protected" accounts, this information is publicly accessible. Table 65 provides a breakdown of the responses to this statement, stratified by user-type.

Table 65

Crosstabulation of User-type and IFQ: Protected Account Meta Information Flow (Q44)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	12	13	84	109
% within group	11.0%	11.9%	77.1%	
Adjusted std. residual	(- 2.7)	(- 2.8)	(+ 4.6)	
Unregistered user	13	16	33	62
% within group	21.0%	25.8%	53.2%	
Adjusted std. residual	(+ .2)	(+ .9)	(- .9)	
Registered user	58	58	117	233
% within group	24.9%	24.9%	50.2%	
Adjusted std. residual	(+ 2.8)	(+ 1.8)	(- 3.8)	
Formerly reg. non-user	3	6	18	27
% within group	11.1%	22.2%	66.7%	
Adjusted std. residual	(- 1.2)	(+ .1)	(+ .9)	
Totals	86	93	252	431

Note. The statement for this question is incorrect.

A majority of respondents in each user group that they did not know whether this statement is correct or incorrect. Non-users (n = 109) had the highest rate of selecting "don't know" at a rate of 77.1%, and registered users (n = 233) had the lowest at 50.2%.

Within the registered user group, an equal amount of respondents indicated incorrectly that they believe the statement to be accurate as accurately identified the statement as incorrect (24.9% for each group). These findings suggest that there may be a good deal of uncertainty and incorrect understanding of what information remains public even when a user has a "protected account" among registered users. This also suggests that what exactly is "protected" when a user protects their account on Twitter may not be fully or accurately known among this group. A chi-squared test did reveal a statistically significant relationship between user-type and response, $X^2(6, N = 431) = 24.483$, $p < .05$.

Knowledge of Ownership among Sample

Following the questions that dealt with information flows in relation to users on Twitter, participants were brought to the last set of questions regarding information flows on Twitter, questions about the ownership of the site. Based on the information flow framework developed in Chapter 4, ownership is part of the socio-economic formation that influences and shapes the arrangement of information flows on Twitter. For this set of questions, respondents were again instructed that they would be given a correct or incorrect statement and would then be asked to indicate that "Yes, this [statement] is correct," "No, this is incorrect," or "I don't know the answer to this question."

The first statement respondents were presented with is the correct statement: "Twitter is a publicly traded company on the New York Stock Exchange." Table 66 provides a breakdown of the responses to this statement, stratified by user-type.

Table 66

Crosstabulation of User-type and IFQ: Twitter as a Publicly Traded Company (Q45)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	29	5	75	109
% within group	26.6%	4.6%	68.8%	
Adjusted std. residual	(- 2.3)	(- .6)	(+ 2.6)	
Unregistered user	22	3	36	61
% within group	36.1%	4.9%	59.0%	
Adjusted std. residual	(+ .0)	(- .3)	(+ .1)	
Registered user	93	15	127	235
% within group	39.6%	6.4%	54.0%	
Adjusted std. residual	(+ 1.7)	(+ .6)	(- 2.0)	
Formerly reg. non-user	11	2	14	27
% within group	40.7%	7.4%	51.9%	
Adjusted std. residual	(+ .5)	(+ .4)	(- .7)	
Totals	155	25	252	432

Note. The statement for this question is correct.

More than half of each user group indicated that they did not know whether this statement was correct or incorrect, with non-users (n = 109) having the highest "don't know" response rate at 68.8% and formerly registered non-users (n = 27) having the lowest at 51.9%. Formerly registered non-users had the highest rate of accurately identifying this statement as correct at 40.7%, followed by registered users (n = 235) at 39.6%, unregistered users (n = 61) at 36.1%, and non-users at 26.6%. Across all groups, only 25 of 432 respondents (5.7%) inaccurately identified this statement as incorrect. A subsequent chi-squared test showed that there was no statistically significant relationship between user-type and response to this question, $X^2(6, N = 432) = 7.294$, p = .295.

Respondents were then presented with the last information flow prompt, the incorrect statement: "Tim Cook is the current CEO of Twitter." This statement is incorrect as Tim Cook is the CEO of Apple. It is worth noting that this survey concluded one week before Tim Cook publicly stated that he is gay (Cook, 2014), and, as a result,

his name may have been less recognized at the time than it would have been a week later. Table 67 provides a breakdown of the responses to this statement.

Table 67

Crosstabulation of User-type and IFQ: CEO of Twitter (Q46)

Status	Statement is Correct	Statement is Incorrect	Don't Know	Totals
Non-user	5	11	93	109
% within group	4.6%	10.1%	85.3%	
Adjusted std. residual	(- .3)	(- 1.6)	(+ 1.6)	
Unregistered user	2	12	48	62
% within group	3.2%	19.4%	77.4%	
Adjusted std. residual	(- .7)	(+ 1.1)	(- .6)	
Registered user	14	39	182	235
% within group	6.0%	16.6%	77.4%	
Adjusted std. residual	(+ .9)	(+ 1.2)	(- 1.5)	
Formerly reg. non-user	1	2	24	27
% within group	3.7%	7.4%	88.9%	
Adjusted std. residual	(- .3)	(- 1.1)	(+ 1.2)	
Totals	22	64	347	433

Note. The statement for this question is incorrect.

More than three quarters of each user group indicated that they did not know the answer to this question. Unregistered users (n = 62) had the highest rate of accurately identifying this statement as incorrect (19.4%), followed by registered users (n = 235) at 16.6%, non-users (n = 109) at 10.1% and formerly registered non-users (n = 27) at 7.4%. Overall, the number of respondents who inaccurately identified this statement as correct was fairly low, only 22 of 433 (5.1%). A subsequent chi-squared test showed no statistically significant relationship between user-type and response, $X^2(6, N = 433) = 5.837$, p = .442. Knowledge about the leaders of Twitter may not be widespread, but appears fairly uniform across different user-type groups.

Conclusion: Accurate, Inaccurate, and Uncertain Knowledge of Information Flows

There are a number of conclusions that can be drawn from this analysis. First, broadly speaking, the distribution of responses was frequently related with user-type. Non-users, on the whole, were more likely to indicate that they did not know something

than unregistered users, formerly registered non-users, and registered users. There were a few cases, however, where the answer distribution patterns were not statistically related with user-status. Those moments represent important findings because they suggest, for those questions, registered users, formerly registered non-users, unregistered users, and non-users all demonstrate a fairly equal distribution of accurate, inaccurate, and uncertain knowledge. To put this another way, one could say that, for those particular questions, being a registered user is related with having no better knowledge (or worse knowledge) in terms of a statistically significant relationship, than a non-user, unregistered user, or formerly registered non-user. If registered users (who may have been exposed to Twitter, Inc.'s discourse and to the technology itself) fare no better or no worse when asked about a particular facet of an information flow than non-users (who have not been exposed to these materials), this suggests that careful attention should be given to Twitter's Inc.'s discourse around that particular facet. This "no better knowledge" situation occurred in the questions/prompts related to: whether Tweets are ever withheld in countries; whether information about third-party website browsing is sent back to Twitter via buttons and widgets; the fact that Twitter generates revenue by selling access to the full stream of real-time Tweets created by users to third-parties; what the rules are that govern what users of the APIs can and cannot do with Twitter Content; whether or not Twitter's Certified Products partners are allowed to resell Tweets; who the leaders of Twitter are; and finally, the fact that Twitter is a publicly traded company. It is note-worthy that these facets commonly have to do with the third-party data sharing, revenue generation based on user-content creation, governance, and ownership.

Second, as seen in Table 68, a rank ordering the rates of accurate, inaccurate, and don't know responses helps produce a picture of what facets of the information flows on Twitter are well understood by registered users, misunderstood by registered users, and where registered users have uncertainty in their *principles-knowledge*.

Table 68

Rank Ordered Chart of Registered Users' Accurate, Inaccurate, and "Don't Know" Responses

	Accurate Identification	Inaccurate Identification	Don't Know
1	How Hashtags Work (Q11) (94.9%)	Protected @Replies (Q15) (51.5%)	Do Certified Products Resell Tweets (Q40) (80.2%)
2	How Following Works (Q14) (92.7%)	@Replies and Timelines. (Q19) (45.1%)	CEO of Twitter (Q46) (77.4%)
3	Revenue Generation, Promoted Tweets (Q31-O1) (73.2%)	Direct Message Defaults (Q20) (44.7%)	Tweets Ever Withheld (Q8) (73.5%)
4	What Makes a Tweet a @Reply (Q12) (72.3%)	Protected Tweets Visible to Follower-Followers (Q9) (35.7%)	Governance of Twitter API Content Rebroadcasting (Q34) (71.9%)
5	Protected Accounts and Following Approval (Q41) (72.2%)	Unregistered User Access to Search (Q38) (33.3%)	Location-based Search API (Q28) (71.6%)
6	Verified Account Authentication (Q42) (70.6%)	Revenue Generation, Buttons and Widgets (Q31-O4) (31.9%)	Twitter Receives Data about Site Visits from Buttons and Widgets (Q27) (67.8%)
7	Tweet Length (Q6) (64.7%)	@Replies Visibility on Followers' Timelines. (Q18) (25.4%)	Twitter Offers APIs (Q25) (67.4%)
8	Favorites Accessible to Others (Besides Author) (Q13) (58.1%)	Protected Account Meta Information Flow (Q44) (24.9%)	Tweets Archived by LoC (Q37) (67.1%)
9	Revenue Generation, Promoted Trends (Q31-O5) (56.2%)	Defaults and Geo-location on Tweets (Q23) (22.6%)	Twitter Warns about Possibly Sensitive Tweets (Q10) (65.8%)
10	Immediate Popularity of Trending Topics (Q16) (51.9%)	Revenue Generation, "Tweet This" Clicks (Q31-O8) (17.9%)	Does Twitter Delete Old Tweets (Q26) (62%)
11	Protected Account Defaults (Q21) (51.3%)	Twitter Receive Data about Site Visits from Buttons and Widgets (Q27) (16.7%)	Governance of Spam Content (Q33) (57.1%)
12	Buttons and Widgets Leading to URL (Q30) (51.1%)	Revenue Generation, Media Charges (Q31-O6) (15.7%)	Twitter as a Publicly Traded Company (Q45) (54%)
13	Users Can Share GPS Data in Tweets (Q7) (51.1%)	Protected Account Defaults (Q21) (15.4%)	Twitter's Three Governance Documents (Q32) (53.6%)
14	Default Tailoring People Suggested (Q22) (50.4%)	Twitter Warns about Possibly Sensitive Tweets (Q10) (15.4%)	Defaults and Geo-location on Tweets (Q23) (53.4%)
15	Info Advertisers Receive (Q43) (49.4%)	Tweet Length (Q6) (14.5%)	Defaults and Tailored Advertisements (Q24) (52.8%)
16	Governance of Tweet Content (Q36) (48.7%)	What Trending Topics Users See (Q17) (14.5%)	Retrieving Old Tweets (Q29) (51.1%)
17	Revenue Generation, Promoted Accounts (Q31-O7) (47.2%)	Are Verified Accounts Paid For (Q39) (14.5%)	Protected Account Meta Information Flow (Q44) (50.2%)
18	What Trending Topics Users See (Q17) (45.3%)	Favorites Accessible to Others (Besides Author) (Q13) (13.2%)	Info Advertisers Receive (Q43) (50.2%)
19	Retrieving Old Tweets (Q29) (44.6%)	Tweets Archived by LoC (Q37) (11.5%)	@Replies Visibility on Followers' Timelines. (Q18) (50%)

	Accurate Identification	Inaccurate Identification	Don't Know
20	Defaults and Tailored Advertisements (Q24) (44.3%)	What Makes a Tweet a @Reply (Q12) (11.1%)	Are Verified Accounts Paid For (Q39) (46.8%)
21	Twitter as a Publicly Traded Company (Q45) (39.6%)	Revenue Generation, Verified Accounts (Q31-O2) (10.6%)	Unregistered User Access to Search (Q38) (46.2%)
22	Are Verified Accounts Paid For (Q39) (38.7%)	Twitter's Three Governance Documents (Q32) (10.6%)	Governance of Tweet Content (Q36) (46.2%)
23	Twitter's Three Governance Documents (Q32) (35.7%)	Do Certified Products Resell Tweets (Q40) (10.3%)	Buttons and Widgets Leading to URL (Q30) (45.5%)
24	Governance of Spam Content (Q33) (35.6%)	Users Can Share GPS Data in Tweets (Q7) (9.9%)	Default Tailoring People Suggested (Q22) (45.3%)
25	Does Twitter Delete Old Tweets (Q26) (34.6%)	Tweets Ever Withheld (Q8) (8.5%)	Immediate Popularity of Trending Topics (Q16) (41.7%)
26	Twitter Offers APIs (Q25) (31.3%)	Governance of Spam Content (Q33) (7.3%)	@Replies and Timelines. (Q19) (40.8%)
27	Protected Tweets Visible to Follower-Followers (Q9) (30.6%)	Immediate Popularity of Trending Topics (Q16) (6.4%)	Protected @Replies (Q15) (40%)
28	Location-based Search API (Q28) (25.9%)	Twitter as a Publicly Traded Company (Q45) (6.4%)	What Trending Topics Users See (Q17) (39.7%)
29	Governance of Twitter API Content Rebroadcasting (Q34) (25.5%)	CEO of Twitter (Q46) (6%)	Users Can Share GPS Data in Tweets (Q7) (39.1%)
30	Protected Account Meta Information Flow (Q44) (24.9%)	Governance of Tweet Content (Q36) (5.1%)	Direct Message Defaults (Q20) (34.9%)
31	@Replies Visibility on Followers' Timelines. (Q18) (24.6%)	Default Tailoring People Suggested (Q22) (4.3%)	Protected Tweets Visible to Follower-Followers (Q9) (33.6%)
32	Defaults and Geo-location on Tweets (Q23) (23.9%)	Retrieving Old Tweets (Q29) (4.3%)	Protected Account Defaults (Q21) (33.3%)
33	Revenue Generation, Firehose Access (Q31-O3) (21.7%)	Buttons and Widgets Leading to URL (Q30) (3.4%)	Revenue Generation, Do Not Know (Q31-O9) (31.1%)
34	Tweets Archived by LoC (Q37) (21.4%)	Does Twitter Delete Old Tweets (Q26) (3.4%)	Favorites Accessible to Others (Besides Author) (Q13) (28.6%)
35	Unregistered User Access to Search (Q38) (20.5%)	Defaults and Tailored Advertisements (Q24) (3%)	Verified Account Authentication (Q42) (27.3%)
36	Direct Message Defaults (Q20) (20.4%)	Location-based Search API (Q28) (2.6%)	Protected Accounts and Following Approval (Q41) (25.6%)
37	Twitter Warns about Possibly Sensitive Tweets (Q10) (18.8%)	Governance of Twitter API Content Rebroadcasting (Q34) (2.6%)	Tweet Length (Q6) (20.9%)
38	Tweets Ever Withheld (Q8) (17.9%)	Verified Account Authentication (Q42) (2.2%)	What Makes a Tweet a @Reply (Q12) (16.6%)
39	CEO of Twitter (Q46) (16.6%)	Protected Accounts and Following Approval (Q41) (2.1%)	How Following Works (Q14) (6.4%)
40	Twitter Receive Data about Site Visits from Buttons and Widgets (Q27) (15.5%)	How Hashtags Work (Q11) (1.7%)	How Hashtags Work (Q11) (3.4%)
41	@Replies and Timelines. (Q19) (14.2%)	Twitter Offers APIs (Q25) (1.3%)	
42	Do Certified Products Resell Tweets (Q40) (9.5%)	How Following Works (Q14) (0.9%)	
43	Protected @Replies (Q15) (8.5%)	Info Advertisers Receive (Q43) (0.4%)	

Note. The true, false, and "don't know" revenue generation methods are only listed once in each column with no matching sets.

This rank ordering suggests a sort of topology of the concentrated areas in which registered users have accurate, inaccurate, and uncertain knowledge of information flows. Further, this rank ordering also suggests specific aspects of Twitter, Inc.'s discourse about information flows that should garner careful and close examination. For example, hashtags, following, and the fact that Twitter uses promoted Tweets to generate revenue are fairly well understood by registered users. As a result, one might ask if there is something in the language that Twitter presents to users about these aspects that is particular clear or if Twitter gives an abundance of attention to these elements in its language for users. Conversely, there is a higher degree of inaccurate knowledge about protected @Replies, the defaults for direct messaging, and the visibility of protected Tweets to the followers of a user's followers. Is there something in Twitter's discourse that might suggest why this is? Are these elements explained less clearly or less frequently in Twitter's language? Finally, registered users displayed a high degree of uncertainty about (for example) whether or not participants in Twitter's Certified Products program resell Tweets, whether or not Tweets are even withheld on a country-by-country basis, whether Twitter receives data about third-party site visits via buttons and widgets, whether or not Twitter deletes old Tweets, and the fact that Tweets are archived by the Library of Congress. How does Twitter's discourse address these facets? Do they address these facets? How frequently and in what way?

Third, in looking holistically across the findings from this analysis, a snapshot of some of the potential limits of users' informational power begins to emerge. For example, based on the responses to the questions about how Twitter generates revenue, it appears that registered users are quick to recognize the ways that promoted accounts, promoted

trends, and promoted Tweets function as revenue generation for Twitter. Yet, the fact that Twitter sells access to user-generated content via the firehose is fairly uncertain among registered users. This suggests that users may conceptualize Twitter as supported through advertising, but not inherently through the sale of access to user-generated content. This does not fully match the picture of the business practices as part of the socio-economic dimension of information flows from Chapter 4. But where did this perception come from? Is it possible that Twitter's own discourse would promote such an interpretation? Is it possible that, as Scholz (2008) might suggest, Twitter, Inc. would seek to limit such *principles-knowledge* in order to keep from alienating their user/labor base? Additionally, registered users displayed a good deal of uncertainty about whether or not the Library of Congress archives Tweets, whether Twitter deletes old Tweets, and whether or not Twitter's Certified products are allowed to resell historical data. This suggests that the flow of older information may be unclear to some registered users. Again, this raises the questions about the antecedents to such uncertainty.

Lastly, "don't know" responses were selected more frequently than inaccurate responses among registered users. This suggests that outright misunderstanding of the facets of information flows on Twitter may not be as prevalent as uncertainty in registered users' *principles-knowledge* of the facets of information flow. While scholars such as Fuchs (2009) have deployed surveys that ask respondents to identify whether particular statements about surveillance practices on social media site are correct or incorrect, the results of the analysis in this chapter suggest that some nuance may be lost when respondents are left to make guesses when they are instead uncertain. These findings suggest that it may be more methodologically beneficial to measure for

uncertainty and ambiguity in user knowledge, at least in exploring understandings of social media. This also raises an important consideration then looking ahead to the discourse analysis: how is ambiguity or uncertainty produced or maintained through discourse.

Chapter 7: Discourse Analysis of User Orientation

Introduction

This chapter addresses the second operationalized research question of this project (RQ2): How does the technological discourse surrounding the site created by Twitter's business purveyors describe information flows? As detailed in Chapter 5, this project uses the method of critical discourse analysis to explore how language presented to users in the sign up process depicts information flows on the Twitter platform, with an eye towards the implications these depictions may have for the development of users' *principles-knowledge* of the facets of information flow, a key component of an individual's informational power. This chapter begins by first providing a general description of the documents that constitute the corpus. Next, as Fairclough (1995) suggests critical discourse analysis should account for the context of discourse in addition to the content, this chapter turns to describe some of the contextual factors that played into the production of these documents. Following this, the chapter outlines the trends that emerged from the application of the coding tool to the corpus. The chapter then returns to comment on a handful of aspects of the corpus that fell outside the coding scheme which may additionally have consequences for users' *principles-knowledge.* Finally, this chapter concludes by summarizing how the context and content of this discourse may influence informational power in the Twitter-user relationship.

Description of the Text Corpus

The corpus of text under consideration in this chapter includes: 1) Twitter's "landing page" of Twitter.com, 2) Twitter's "Join Twitter Today" page; 3) Twitter's Terms of Service (TOS), Privacy Policy, the Twitter Rules, and Twitter's Cookie Use

statement; 4) Twitter's new user orientation tutorial; and 5) the web interface that newly registered users are brought to once they have completed the sign up process. While each of these documents are unique entities, some of which engage different speech genres, for the purposes of this dissertation they are being considered as a coherent body of text. This is because they are the documents that a user encounters in the process of coming to Twitter.com and registering for an account.

The sign-up process represents a key moment for Twitter, Inc. to convince individuals who are new to the platform to adopt the technology, to become registered users, and to orient them to the platform. Both growing its user-base and orienting users towards information production are activities critical for Twitter, Inc.'s revenue generation. As a result, in these texts, one might expect to see Twitter, Inc. offering arguments about the advantages of its product in order to persuade users towards adoption. One might also expect to see language that would promote *how-to knowledge* about the platform, but not necessarily detailed *principles-knowledge* of information flows on the platform, particularly if that *principles-knowledge* relates to information flows that might give users a reason to abandon the registration process or would give them reason to hesitate in the production of content. Conversely, this process is also important for individuals contemplating the adoption of the technology. It is particularly important for individuals who may not yet have a widely developed set of *how-to knowledge* or *principles-knowledge* about the Twitter platform, but who have enough *awareness-knowledge* to visit the Twitter.com webpage. It is a chance for them to develop or reinforce their own knowledge about of the facets of information flow on the site.

For some individuals, it may be the first time they are encountering the Twitter.com website. They may be starting from scratch in their *how-to* and *principles-knowledge* of Twitter, and as a result, the language they encounter may play a significant role in the construction of their mental models of information flows on the platform. Other individuals may already be registered users who are simply interested in creating a second, third, or hundredth account. In that case, the language they encounter may reinforce their already established *awareness*, *how-to,* or *principles knowledge*. In either case, the sign-up process represents an important moment in the Twitter-user relationship in which informational power is negotiated vis-à-vis the consumption of discourse about information flows on the service.

Screen-captures and HTML text-only captures of the documents under analysis were saved from the Twitter.com homepage on October 31st, 2013. The *Context: History and Production of the Text Corpus* section of this chapter provides explanation about why this particular moment is unique in Twitter's history and why it has bearing on the content of the corpus to some degree.

Landing Page.

The Twitter.com landing page is the first document an individual signing up for a registered account on Twitter would encounter. This "welcome page" is a single page screen. A horizontal black bar at the top of the webpage contains a small Twitter logo and, in the top right, a drop down menu allows a user to select the language presented on the page. Only the English language version of the site was captured for analysis. Below the black bar at the top, there are three white rectangular boxes. The left most white box welcomes the reader to Twitter. The top right box allows already registered users to sign-

in, and the bottom right box allows a user who is new to Twitter to begin the sign up process by entering their "Full Name, "E-mail," "Password," and clicking the "Sign up for Twitter" button. This can be seen in Figure 17 below.

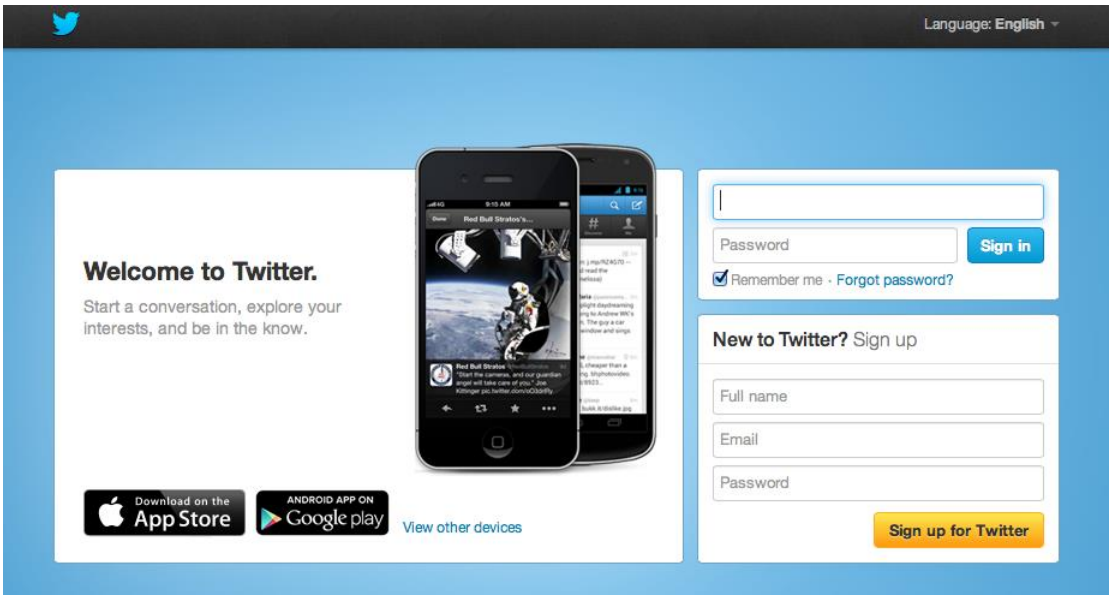


Figure 17. Twitter.com Landing Page

Join Twitter Today.

Once a user has clicked on the "Sign up for Twitter" button on the landing page, they are brought to the second document in the corpus, the "Join Twitter Today" page. On this page users can fill in a full name, an e-mail address, a password, and choose a user-name. After these inputs, there are two selection boxes that give the user the option to stay signed in on a particular computer (via a session cookie) and to "Tailor Twitter based on my recent website visits" (which is selected by default, unless a user has the "Do Not Track" option activated within their web-browser). A text box appears below this, with the first two lines of Twitter's "Terms of Service" visible in the box. If a user clicks on the text-box, it expands, making more lines of the ToS visible. Links to printable versions of the Terms of Service, the Privacy Policy, and the Cookie Use

statement appear to the right of the text-box containing the ToS. Below the expanding text box containing the ToS, a large yellow button states “Create my account.” Below this, in smaller font and in light color than the rest of the text on the page, is a statement regarding the discoverability of Twitter users and some information about privacy settings (described further in the content analysis section of this chapter). This can be seen in Figure 18 below.

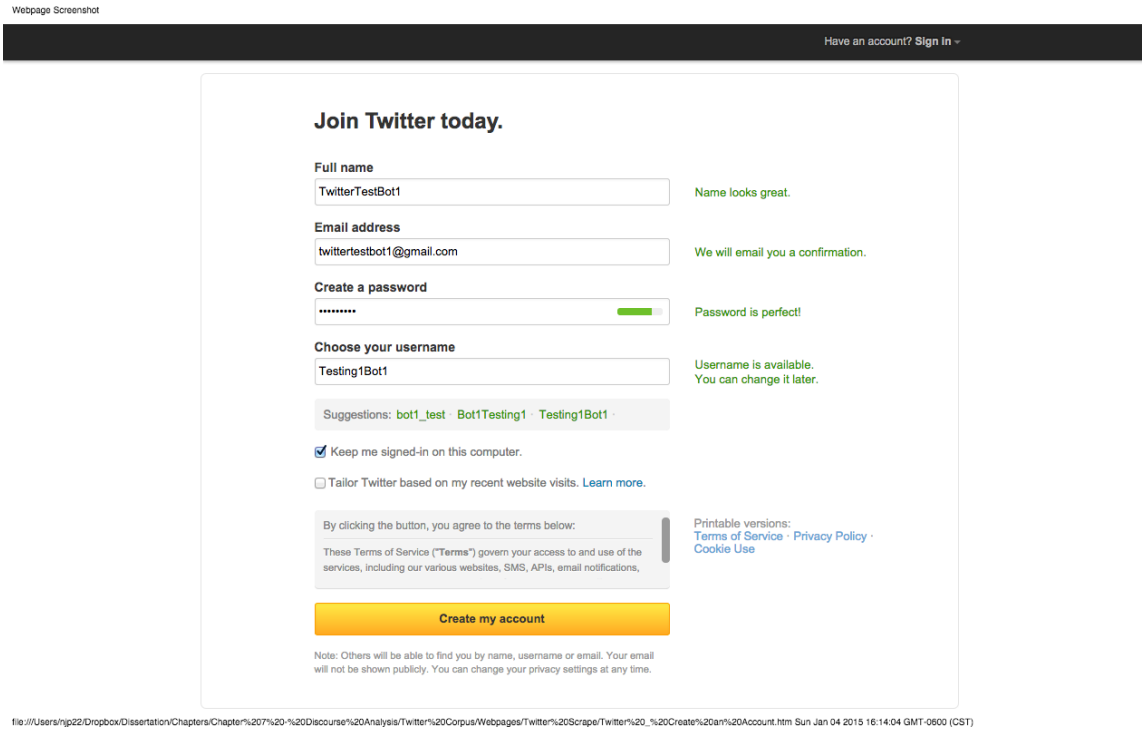


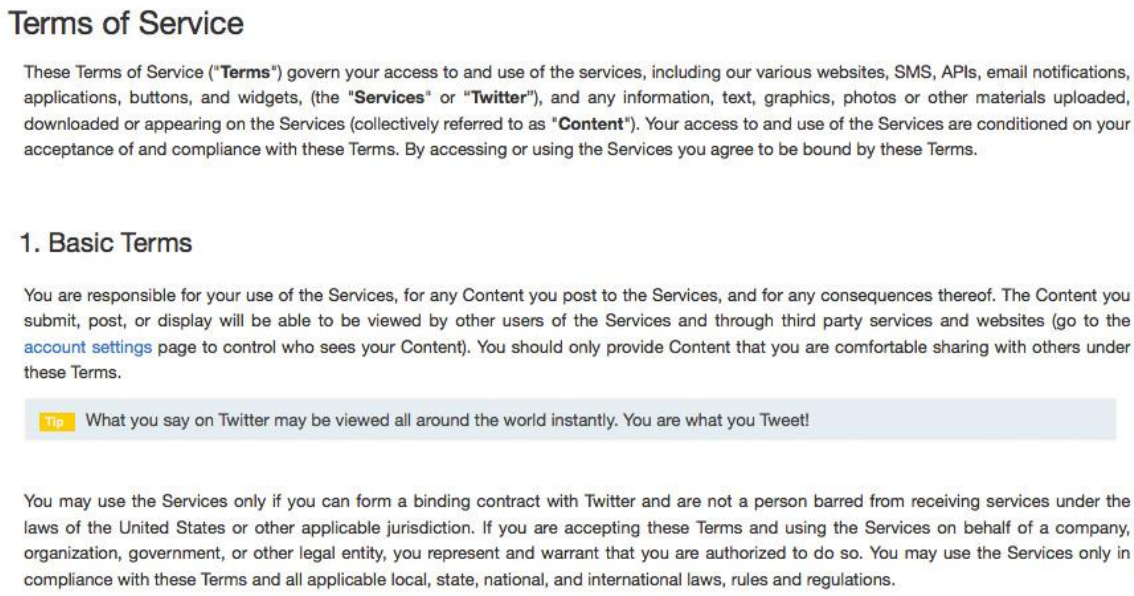
Figure 18. The “Join Twitter today” page.

Policy Documents.

Twitter’s Terms of Service, Privacy Policy, and Twitter’s Cookie Use statement are all hyperlinked within the “Join Twitter Today” page. However, the Terms of Service notes that one additional document, “The Twitter Rules,” is also considered to be part of the user-agreement. Together, these four documents will be referred to as the “policy documents.” Despite the fact that the Terms of Service technically appears within the

"Join Twitter Today" page, for the purposes of this dissertation, it is being treated as a separate document. It should be noted that users are only confronted with the first two lines of the Terms of Service as part of the sign-up process, and they must actively choose to either scroll down in the textbox on the "Join Twitter Today" page or click on the links to the printable versions of the pages in order to see their full content. Twitter's policy documents are generally longer, more text heavy documents than the other documents described so far.

Twitter's printable "Terms of Service" link leads to "Version 7" of the document, which was created June 25, 2012. The document is 3,482 words long and contains nearly no images. The document contains 12 numbered sections of text, many of which are broken out into subsections. The sections are titled with the headers: "1. Basic Terms," "2. Privacy," "3. Passwords," "4. Content on the Services," "5. Your Rights," "6. Your License to Use the Services," "7. Twitter Rights," "8. Restrictions on Content and Use of the Services," "9. Copyright Policy," "10. Ending These Terms," "11. Disclaimers and Limitations of Liability," and "12. General Terms" As shown in Figure 19, the text is occasionally punctuated by "Tips" throughout the document.



Terms of Service

These Terms of Service ("**Terms**") govern your access to and use of the services, including our various websites, SMS, APIs, email notifications, applications, buttons, and widgets, (the "**Services**" or "**Twitter**"), and any information, text, graphics, photos or other materials uploaded, downloaded or appearing on the Services (collectively referred to as "**Content**"). Your access to and use of the Services are conditioned on your acceptance of and compliance with these Terms. By accessing or using the Services you agree to be bound by these Terms.

1. Basic Terms

You are responsible for your use of the Services, for any Content you post to the Services, and for any consequences thereof. The Content you submit, post, or display will be able to be viewed by other users of the Services and through third party services and websites (go to the account settings page to control who sees your Content). You should only provide Content that you are comfortable sharing with others under these Terms.

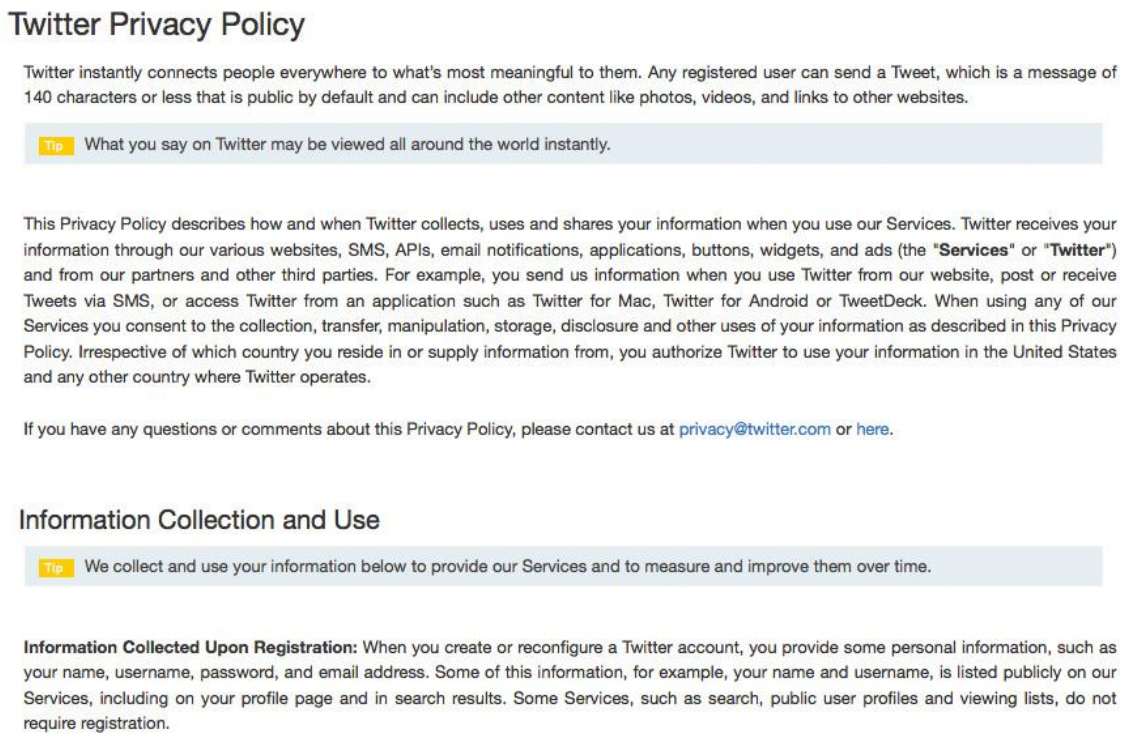
Tip What you say on Twitter may be viewed all around the world instantly. You are what you Tweet!

You may use the Services only if you can form a binding contract with Twitter and are not a person barred from receiving services under the laws of the United States or other applicable jurisdiction. If you are accepting these Terms and using the Services on behalf of a company, organization, government, or other legal entity, you represent and warrant that you are authorized to do so. You may use the Services only in compliance with these Terms and all applicable local, state, national, and international laws, rules and regulations.

Figure 19. Screencapture of a portion of Twitter's Terms of Service including "Tip."

The document is also interspersed with hyperlinks to other pages, such as links to the Twitter account settings page, the Notifications settings page, the Twitter Rules, Twitter's Privacy Policy, and Twitter's API developer's page, among others.

Twitter's "Privacy Policy" is "Version 8" of the document. It was created October 21, 2013 and is 2,266 words long. Unlike the Terms of Service, the Privacy Policy does not include a numbered set of headings, though it does contain various headings such as "Information Collection and Use," "Information Sharing and Disclosure," "Modifying Your Personal Information," "Our Policy Towards Children," "EU Safe Harbor Framework," and "Changes to this Policy." Similar to the Terms of Service, the document also includes a set of "tips" interspersed throughout the document, as seen in Figure 20.



Twitter Privacy Policy

Twitter instantly connects people everywhere to what's most meaningful to them. Any registered user can send a Tweet, which is a message of 140 characters or less that is public by default and can include other content like photos, videos, and links to other websites.

Tip What you say on Twitter may be viewed all around the world instantly.

This Privacy Policy describes how and when Twitter collects, uses and shares your information when you use our Services. Twitter receives your information through our various websites, SMS, APIs, email notifications, applications, buttons, widgets, and ads (the "**Services**" or "**Twitter**") and from our partners and other third parties. For example, you send us information when you use Twitter from our website, post or receive Tweets via SMS, or access Twitter from an application such as Twitter for Mac, Twitter for Android or TweetDeck. When using any of our Services you consent to the collection, transfer, manipulation, storage, disclosure and other uses of your information as described in this Privacy Policy. Irrespective of which country you reside in or supply information from, you authorize Twitter to use your information in the United States and any other country where Twitter operates.

If you have any questions or comments about this Privacy Policy, please contact us at privacy@twitter.com or here.

Information Collection and Use

Tip We collect and use your information below to provide our Services and to measure and improve them over time.

Information Collected Upon Registration: When you create or reconfigure a Twitter account, you provide some personal information, such as your name, username, password, and email address. Some of this information, for example, your name and username, is listed publicly on our Services, including on your profile page and in search results. Some Services, such as search, public user profiles and viewing lists, do not require registration.

Figure 20. Screencapture of a portion of Twitter's Privacy Policy including "Tip."

The document has a number of hyperlinks to other pages, such as links to the account setting pages and support pages for parents, among others.

The "Twitter Rules" do not have a version number, nor an explicit date created, although a copyright notice does indicate the text is copyright 2013. The Twitter Rules are 1,213 words long. Whereas the Terms of Service and Privacy Policies are standalone documents, the Twitter Rules appear as a subpage within Twitter's broader "Help Center." Other than the Cookie Use document and The Twitter Rules, no other parts of the Help Center were included in the corpus. As can be seen in Figure 21, unlike the Terms of Service or Privacy Policies, there are links at the top of the page to "Welcome to Twitter," "Me," "Connect," "Discover," "Mobile & Apps," and "Troubleshooting," and links on the left hand side of the page leading to pages titled "Understanding your

safety settings," "Choosing your experience," "Dealing with issues online," "Tips," and "Resources."

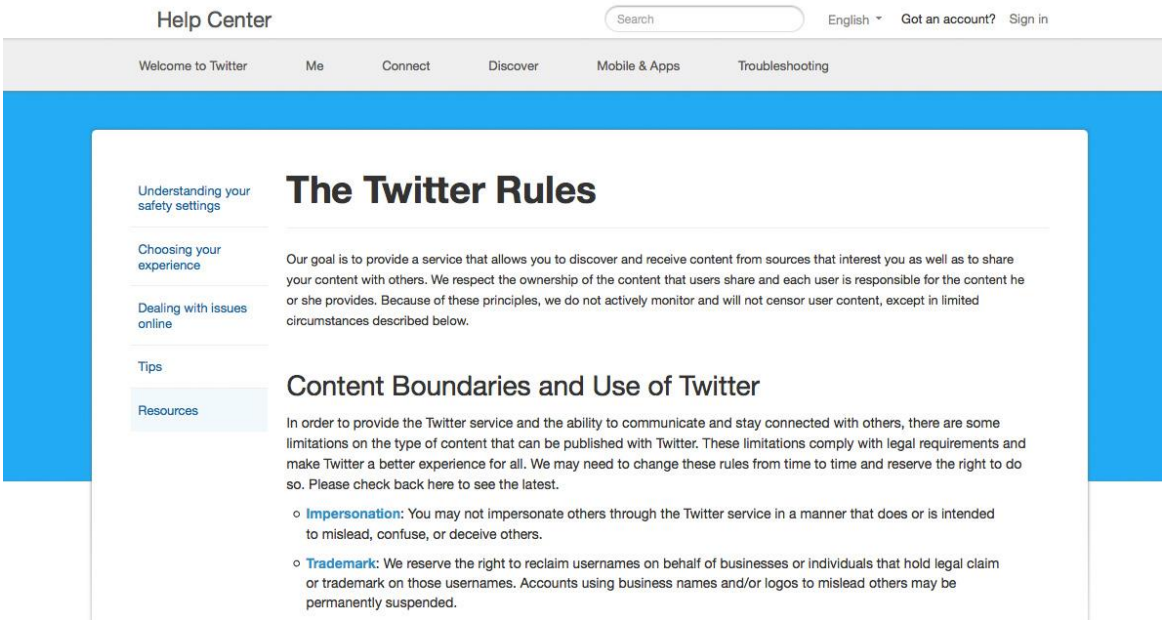


Figure 21. Screencapture of a portion of "The Twitter Rules."

The text of The Twitter Rules contains three sections, "Content Boundaries and Use of Twitter," "Abuse and Spam," and "Have Questions?"

Finally, similar to "The Twitter Rules," the Cookie Use statement is just over a thousand words (1,092), does not have a version number (although there is a 2013 copyright notice), and appears as part of the "Help Center." This document contains five sections, "What are cookies, pixels, and local storage?," "Why does Twitter use these technologies?," "Where are these technologies used?," "What are my privacy options?," and "Where can I learn more?" As can be seen in Figure 22, on the left side of the page, links appear for "Twitter Rules & policies," "Guidelines," "Report a violation," and "Advertiser policies."

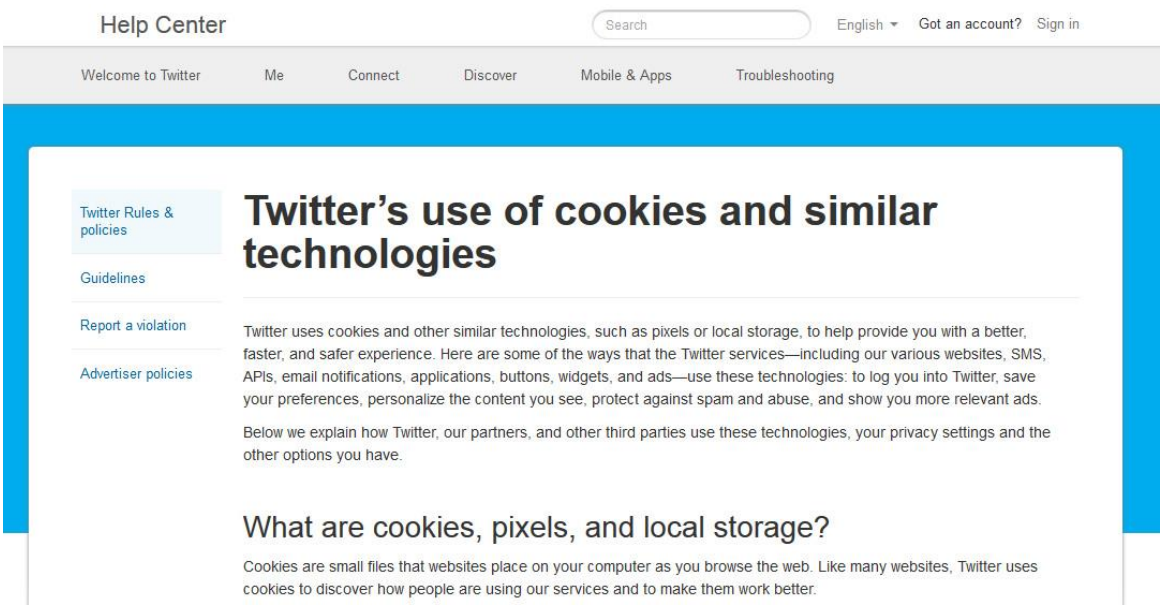


Figure 22. Screencapture of a portion of the Cookie Use statement.

The document has a number of hyperlinks throughout, and the entire final "Where can I learn more?" section is composed of eight hyperlinks to topics such as "Twitter's Privacy Policy," "How Twitter Ads work," "About tailored suggestions," "Twitter supports 'Do Not Track'," "FAQs about tailored suggestions," "Your privacy controls for tailored ads," "FAQs about tailored ads and your privacy preferences," and "Measuring brand impact and your privacy controls." Notably, this is one of the few locations in the corpus where there is such a high concentration of links offering to inform users about different facets of information flows on the platform.

Twitter's new user orientation tutorial.

Once a user has clicked on the "Create my Account" button on the "Join Twitter Today" page, they are brought to a series of pages that this dissertation will refer to as Twitter's "New User Orientation Tutorial." As shown in Figure 23, this process begins with two boxes. One on the left that aids navigation through the tutorial, and one on the right that mimics the timeline interface.

Figure 23. Screencapture of the first step of the new user orientation tutorial.

A newly registered user progresses through the tutorial by hitting the "Next" button. First "The Twitter Teacher" shows users a Tweet and gives a description of what constitutes a Tweet. Next, as can be seen in Figure 24, the newly registered user is told to build their Timeline by choosing five users to follow from a list of suggestions.

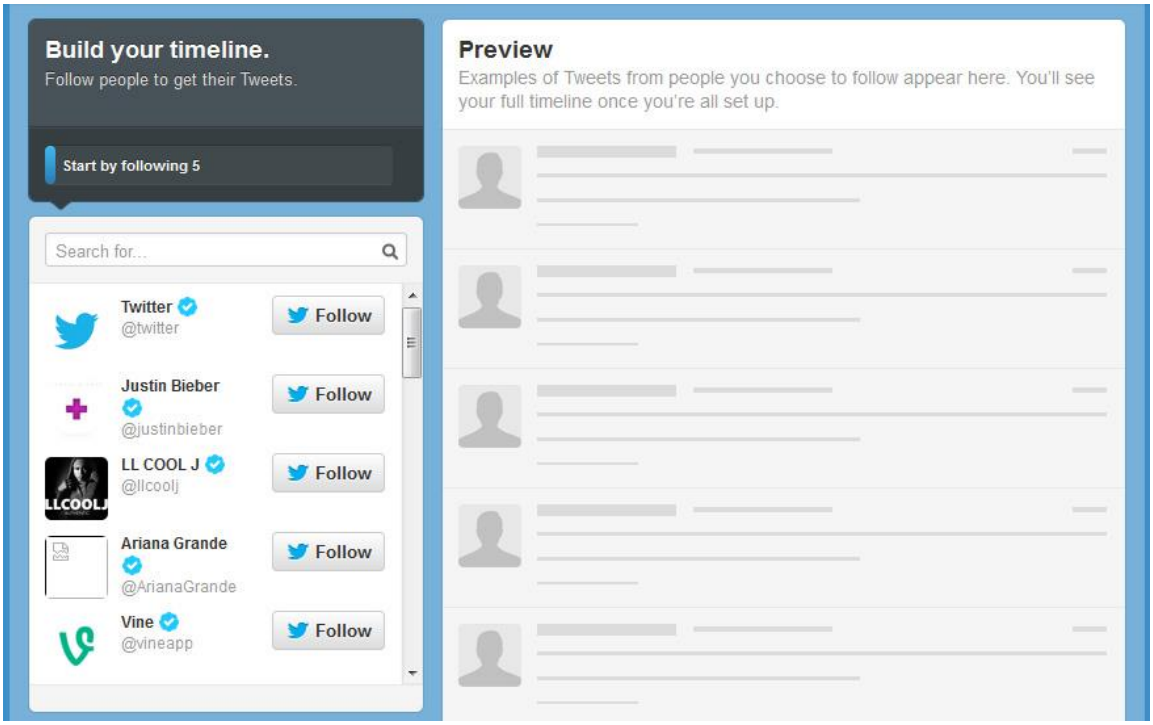


Figure 24. Screencapture of the "Build your timeline" step of the new user orientation tutorial.

As a user clicks on the "Follow" buttons, the most recent Tweets of the user being newly followed are added to the box on the right. Once the registering user has followed five

Twitter users, they are next taken to the "Find People You Know" step as shown in Figure 25.

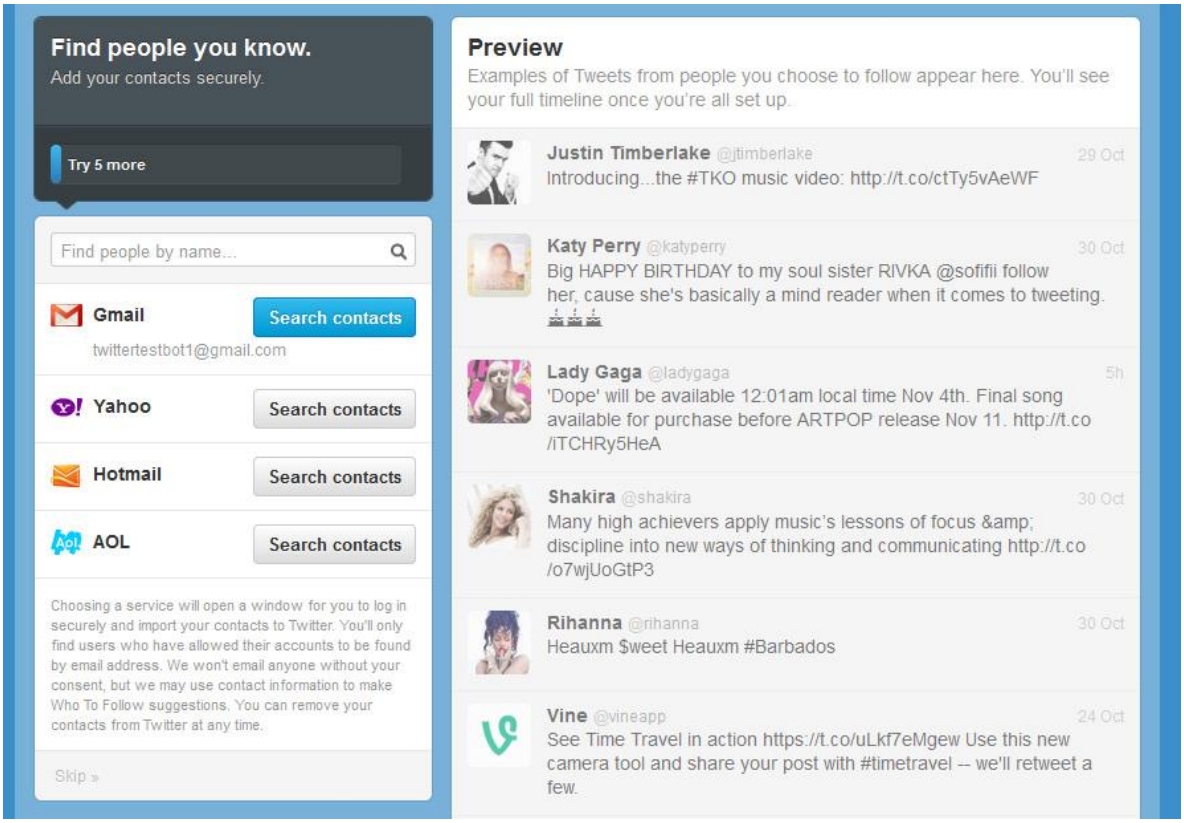


Figure 25. Screencapture of the "Find people you know" step of the new user orientation tutorial.

On this page users are given a search bar, through which they can find other Twitter users by name. This step also allows users of Gmail, Yahoo!, Hotmail, and AOL to import an e-mail address contact list to the platform to automatically search for others. As users follow these contacts, the "Preview" area again populates with the most recent Tweets from the newly added followers.

Finally, the tutorial takes registering users to the "Add character" step. As can be seen in Figure 26, in this step, users are offered the opportunity to add a profile picture and to enter a 160-character biography about themselves.

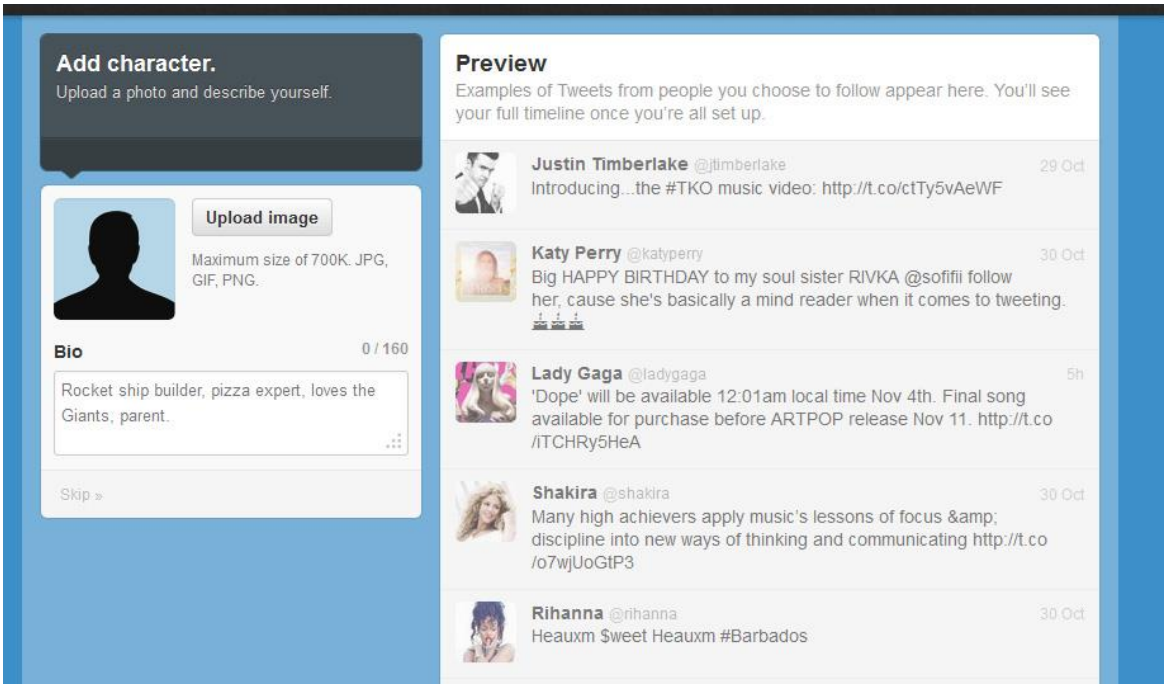


Figure 26. Screencapture of the "Add character" step of the new user orientation tutorial.

Twitter Web-Interface.

After the new user orientation process has been completed, the newly registered user is taken to the web-interface. The web-interface can be seen in Figure 27.

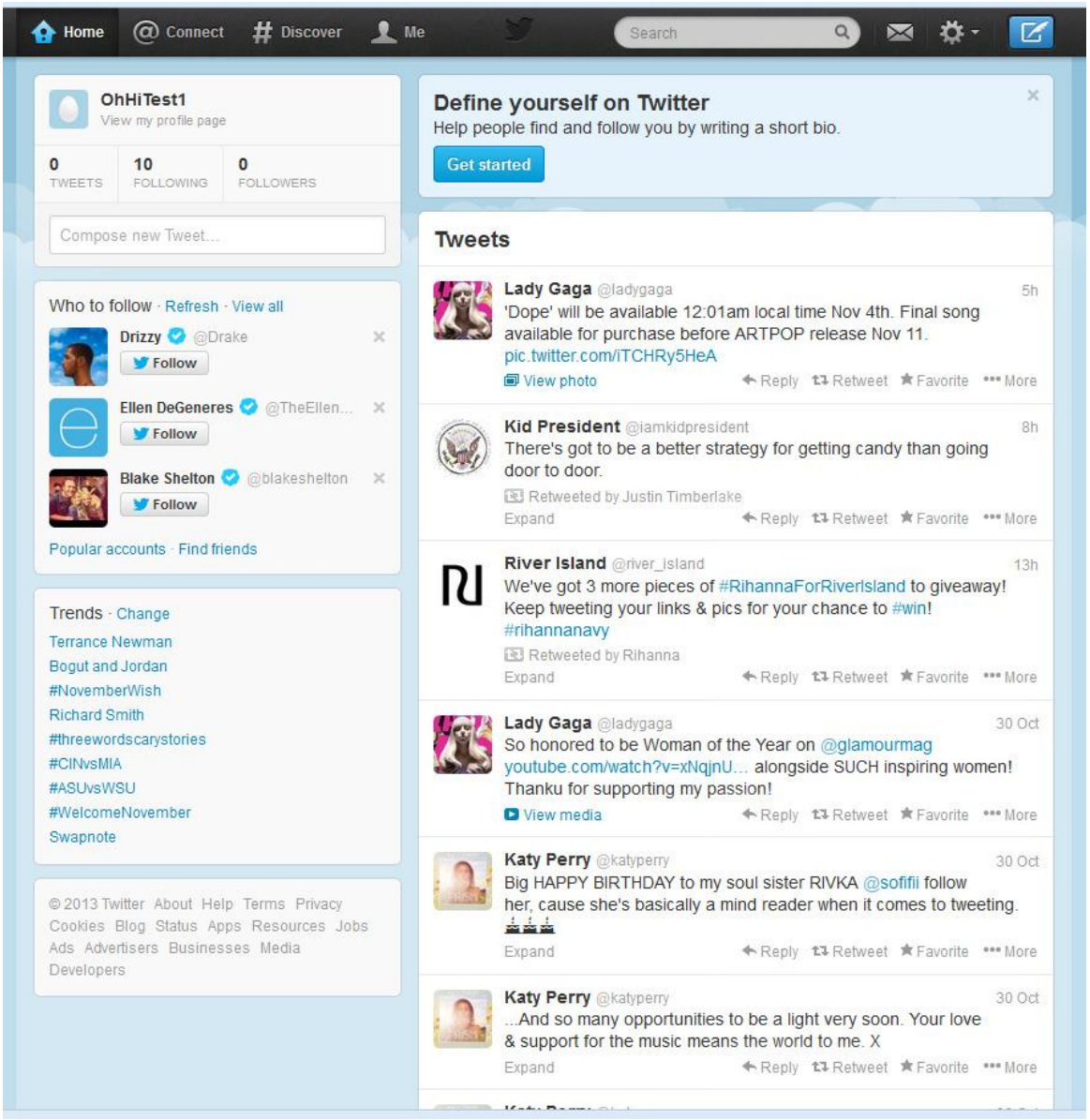


Figure 27. Screencapture of Twitter's web interface.

Once the newly registered user has reached this screen, they are now immersed in the world of Twitter. They can Tweet, interact with their timelines, and can browse the "@Connect," "#Discover," "Me" and settings pages. The "@Connect," "#Discover," "Me" and settings pages are also considered as part of this corpus. As Chapter 4 has already described the web-interface in some detail, this chapter will proceed to discuss the some of the contextual history of these documents.

Context: History and Production of the Text Corpus

Twitter, Inc. produced each of the texts that appear in the corpus at a unique time and place in the history of the platform. Different individuals at Twitter, Inc. were likely involved with the production of the texts, and, as will be discussed further in this section, each was likely produced by Twitter, Inc. with slightly different audiences in mind. The histories of these documents cannot be fully explained absent interviews with the individuals who produced them (which is unfortunately outside the scope of this dissertation), but this chapter will describe a number of contextual factors that are likely to have shaped the production of each document.

Twitter.com landing page.

The Twitter.com landing page has undergone a number of revisions since the site's inception. These changes have included alterations to its design, layout, and content. Each new version of the front-page has included an attempt to communicate to users what Twitter is, and the value that the service offers. In a blog post announcing the site's front-page redesign in 2009, Twitter founder Biz Stone stated:

> Today we're trying a redesigned front page for folks who are new to Twitter.com… Helping people access Twitter in more relevant and useful ways upon first introduction lowers the barrier to accessing the value Twitter has to offer and presents the service more consistently with how it has evolved. (Stone, 2009a, para. 1)

While the front-page is a kind of virtual welcome mat, it is oriented towards an audience wider than just individuals contemplating joining. As stated previously, the version of the front-page included in the corpus was captured in October of 2013. Twitter's initial public offering of stock took place on November 7, 2013. As potential investors may have been visiting the Twitter.com homepage during this time, Twitter,

Inc. may have chosen the language and images that appear on the landing page with these persons additionally in mind.

There is some evidence to suggest this is the case. As reported by the tech-blog *All Things D*, Twitter redesigned its landing page on October 7, 2013, one month before its IPO (Isaac, 2013b). As Isaac (2013) notes:

> The company quietly updated the Twitter.com homepage over the weekend, changing the visual design and welcome text that people see when first encountering the service… the new design is subtle, and focuses on two major points: Mobile, and just exactly what newcomers to Twitter should expect. (paras 2-3)

The foregrounding of mobile access on the front page may not just be for the purposes of communicating to potential *users* that Twitter can be accessed via smartphones, but also to communicate to *investors* who the audience/user base of this product is (mobile users). As, at the time, mobile advertising was seen as an expanding market (eMarketer, 2013), this foregrounding was likely intentionally to assuage investors of any fears they might have about Twitter's potential for revenue growth. Hence, the context of the November IPO like shaped Twitter, Inc.'s choices for the page.

Join Twitter today.

While the landing page serves as a digital welcome mat for more than just individuals contemplating the adoption of the technology, the "Join Twitter Today" page appears much more narrowly focused. The text on this page (absent the ToS which will be described in the section on the policy documents) is fairly minimal and the user only needs to provide input in four places before they can "Create My Account." Users are not required to click into the Terms of Service or proceed to the end of them before they are allowed to register. It appears that this page has been streamlined to put as few hurdles in

the way of the registering user as possible. Given the socio-economic context in which Twitter operates (as a Web 2.0 company whose revenues are impacted by user-base size and use habits/labor), such a minimalist strategy gives readers less text from which to find a reason not to sign up. This economic context likely impacts both the decisions Twitter, Inc. made about words on the page and how those words are presented.

Policy documents.

There are four texts considered as part of the "policy documents": Twitter's Terms of Service, Privacy Policies, the Twitter Rules, and the Cookie Use statement. While the "Join Twitter Today" page was likely written with newly registering users in mind, the policy documents appear to speak to a broader audience. Many of Twitter's policy documents have a distinctly legal tone, a stark contrast to the rather simplistic language present on the Twitter.com landing page and the "Join Twitter Today" page. This is partially attributable to the context in which these documents exist. For example, the language of the Terms of Service and Privacy Policies has likely been influenced by the fact that the Federal Trade Commission (FTC) can take enforcement actions against Twitter, Inc. if it violates the promises of those documents. The Federal Trade Commission's Fair Information Practice Principles of "notice," "choice," "access," "integrity," and "enforcement" also likely inform the construction of the Terms of Service and Privacy Policies documents. Industry norms for what is usually found in a terms of service document, privacy policy statement, community rules, and a cookie use statement also likely influence each individual document.

However, despite being grouped together here under the header of "policy documents," each document also has a unique production context. This section will

briefly comment on the individual situational contexts that may have influenced the language present in these documents.

Terms of Service.

Terms of service are documents that dictate the "rules a person or organization must observe in order to use a service" (PCMag, 2014, p. 1). These documents often cover topics such as, "copyright notices, marketing policies of the respective company, what is the meaning of acceptable user behavior while online, etc." (TermsFeed, 2014). In the U.S., terms of service are considered to be legally binding documents, unless the contract violates state or federal law. If users violate terms of service, they may face use sanctions, have their access revoked by the sites' owners, or may be subject to civil liability, though, with a handful of exceptions, U.S. courts have generally not treated ToS violations as a criminal matter (Chan, 2012). Terms of service help define the relationship between the user and the company, and function as a set of rules regarding what the user can or cannot do on the platform. As a result, terms of service are powerful documents in the negotiation of power between the individual and the sociotechnical system. Frequently, however, these terms are provided in a "take it or leave it" manner. Users are rarely given the opportunity to negotiate the terms and routinely must either agree to the terms as they are written, or must instead choose to not use the service.

As mentioned in the general description of the corpus, the "Terms of Service" under consideration in this analysis is the seventh version of the document. In her analysis of the governance of Twitter, van Dijck (2013) traces the evolution of Twitter's Terms of Service through several iterations, remarking, "When Twitter started in 2006, its terms of service were very general and did not say much about the way users could

deploy the tool to communicate" (p. 83). However, over time this began to change. Further, van Dijck notes that the first version of the ToS was silent about how data from Twitter would be used for advertising of applications, but by 2009, the ToS was updated to explicitly indicate that "all data sent through Twitter may be used by third parties" (p. 84). In 2011, the document was amended to incorporate the introduction of Promoted Tweets and Promoted Trends, and to address the fact that Twitter partnered with third-parties (with Google Analytics being named explicitly) for the purposes of data-mining. Van Dijck summarizes this history of changes to the document by stating: "It is difficult not to read these ToS modifications as the next step in Twitter's new ambition to become an 'information company'" (p. 85).

This lineage of changes to Twitter's ToS sheds a bit of light on the historical context of the document. Further, the text that appears as part of Twitter, Inc.'s Terms of Service is informed not just by Twitter, Inc.'s own desires to govern users and user behavior, but also by the U.S. and regional legal contexts in which it operates, and by the context of the business agreements with third-parties that Twitter has developed. In this way, the text of the document does not just speak to "users," but also speaks to lawyers and courts, government policymakers and regulators, third-party developers, and to Twitter, Inc.'s business partners.

Privacy Policy.

Generally speaking, privacy policies are documents that indicate how an organization collects, uses, and transmits user information. As Reidenberg et al. (2014) argue, privacy policies are extremely important for users: "for all their faults, privacy policies remain the single most important source of information for users to attempt to

learn how companies collect, use, and share data" (p. 1). Despite the importance for users, the language in these texts is often dense, legalistic, and frequently requires college-level reading comprehension (Sherman, 2008). This is partially an outcome of the regulatory environment in which the U.S. based technology companies that write privacy policies exist in.

Privacy policies are often included on web sites that collect and use user-data, but U.S. federal law does not explicitly mandate them. With a few exceptions[29], the U.S. generally relies on companies to self-regulate in posting privacy policies. As Reidenberg et al. (2014) describe it:

> … the United States takes a "notice and choice" approach to Internet privacy. The idea is that companies post their privacy policies, users read and understand policies, and users follow a rational decision-making process to engage with companies offering an acceptable level of privacy. This structure is designed and promoted as a replacement for regulation. (p. 1)

Despite the fact that there is not the explicit requirement for privacy policies, if a U.S. company has one posted, the U.S. Federal Trade Commission may take action against that company if they violate what is stated in the document. The FTC treats violations to a stated privacy policy as a deceptive practice. In 2010, Twitter notably became the first social media company to have an action taken against it by the FTC for violating its own privacy policies (Federal Trade Commission, 2010). The FTC argued (Federal Trade Commission, 2011) that despite the fact that Twitter's Privacy Policy indicated that it would take steps to protect certain non-public user information from being accessible to unauthorized users, that, "for three years from July 2006 to July 2009, Twitter did not

[29] There are a handful of laws that do compel the posting of privacy policies in the U.S., such as the Children's Online Privacy Protection Act, which requires an organization to have a privacy policy if the organization collects information from minors under the age of 13, and the Health Insurance Portability and Accountability Act which requires privacy policies be provided when health care providers collect health information.

take reasonable and appropriate measures to prevent unauthorized administrative control of its system" (IT Law Group, 2014). This resulted in hackers compromising administrator accounts, which were subsequently used to send unauthorized Tweets from the Twitter accounts of Barack Obama and Fox News. Twitter eventually settled the complaint with the FTC, and as a result:

> Twitter is barred for 20 years from misleading its users about the extent to which it protects the security, privacy, and security of non-public consumer information. The agreement requires Twitter to establish, implement, and maintain a comprehensive information security program that is "reasonably designed to protect the security, privacy, confidentiality, and integrity" of nonpublic user information. (IT Law Group, 2014)

Both the regulatory environment in which privacy policies as documents exist and Twitter's previous interactions with the FTC around its own privacy policy likely informs the text that appears within the document. As a result, the text of the document does not just speak to an audience of newly registering users, but also speaks to lawyers and courts, government policymakers and regulators, and specifically to the FTC.

The Twitter Rules.

While they are being treated here as a separate document, Twitter's Terms of Service states that the Twitter Rules are actually part of the Terms of Service (Twitter.com, 2012). Yet, the Twitter Rules do not actually appear inside of the Terms of Service document. Instead, a user must click on the hyperlink under section eight of the ToS, the "Restrictions on Content and Use of the Services" to access the rules. As a result of the inclusion of "The Twitter Rules" as part of the ToS, the Twitter Rules should be considered as having many of the same contextual characteristics as the ToS itself.

The "Twitter Rules," however, have not always been a part of Twitter's Terms of Service. It was not until Version 2 of the ToS, introduced in September of 2009, that

Twitter, Inc. formally included them as part of the ToS. The "Twitter Rules" did exist before this date however, and were used as in Twitter, Inc.'s justification for suspending an account that impersonated the Dalai Lama (Semuels, 2009). Unlike the ToS and the Privacy Policy, Twitter does not offer an archive of older versions of the Twitter Rules, despite the fact that they are changed from time to time. The oldest version of the Twitter Rules that I could locate (via the Internet Archive's Wayback Machine) is dated as having been created on January 14, 2009 (Twitter.com, 2009). While the history of the Twitter Rules likely informs the current arrangement of language on the page, this history is not as lengthy as other documents, such as the Privacy Policy.

Cookie Use statement.

Similar to the Twitter Rules, Twitter does not offer an archive of older versions of the Cookie Use statement, despite the fact that the document may change from time to time. Interestingly, the oldest version of the Cookie Use statement available through the Internet Archive's Wayback Machine is from October 26, 2013, just five days before the version used for the dissertation (Twitter.com, 2013). Screen captures of the "Join Twitter Today" page from 2012 only show links to the Terms of Service and Privacy Policy, suggesting this page (if it existed then) was not linked as part of the new user sign-up process. The lack of availability of older versions makes it difficult to say how long Twitter has had the Cookie Use statement, and as a result, it is difficult to say how historical versions of the text might have played a role in the version used within the corpus.

Twitter's New User Orientation.

Twitter's new user orientation has been redesigned several times since the site's creation. One of the reasons for the redesigns of the orientation process was a concern (as stated by Ev Williams) that "Twitter is too hard" (Williams, 2010b). In 2010, a major redesign took place that introduced "suggested users" and the ability to find people that a new user may already know as part of the sign-up process. A blog entry written by Josh Elman, one of the product designers at Twitter, reflects on this re-design. On Twitter's company blog, he introduced the changes, stating:

> Two of the biggest challenges for new users have been finding accounts to follow that appeal to their interests, and finding their friends and colleagues who tweet. Over time, we've learned that by making suggestions of who to follow, we can help users get going more easily on Twitter. In our new design, we're taking some steps to continue to improve this process. Once a user signs up and selects what they're interested in, we show them some accounts that relate to that interest. Next, we help them find their friends and colleagues by checking their address books, and third we give them a chance to search for anyone we or they missed in this process… We've found that the power of suggestion can be a great thing to help people get started, but it's important that we suggest things relevant to them. (Elman, 2010)

While Twitter's landing page and policy documents appear oriented towards multiple audiences, Twitter's new user orientation appears to be targeted more specifically towards individuals who are unfamiliar with how Twitter works. The comments of Ev Williams and Josh Elman suggest that this text projects an image of how Twitter can and should be used for new users. Hence, the context is not only about building *how-to knowledge*, but also cultivating the desire and interest among these new users to want to return to the technology.

Twitter's Web-Interface.

The Twitter.com user interface has undergone a number of changes since the technology's inception. While some changes have only involved redesigns to the site's

aesthetics, other revampings have involved changes to both the functionality and content of the pages (Twitter.com, 2011). Twitter's web-interface was originally shaped in large part based on the idea of a stream of "status updates" that could be sent via SMS. The idea was based in part on other technologies such as AOL Instant Messaging and blogging platforms such as LiveJournal (as stated by Twitter founders Jack Dorsey and Biz Stone in an interview with Slutsky & Codel, 2006). As users developed practices and norms in relation to the site, Twitter, Inc. responded to many of these by formalizing the practices into new protocols and algorithms, codifying them within the interface. Each newly codified practice led to a slight redesign of the interface, such as when retweets and hashtags were turned into protocols (van Dijck, 2013). However, as the user-interface has undergone these changes, the basic idea of the interface as centered on a timeline of Tweets has remained consistent. This history of previous designs informs the layout of and text that appears as part of the interface pages considered in the corpus.

Content: Textual Analysis of the Corpus

This chapter now to turns to discuss the outcomes of the textual analysis of the corpus. There were a number of trends and themes that emerged after applying the three-tier coding scheme to the corpus. To briefly summarize the coding schema, in the first stage of the coding process, each segment (sentence) of text was examined for whether it mentioned facets of information flow from van Dijck's information flow framework: data/metadata, algorithms, protocols, interfaces, defaults, users (as senders or receivers), types of informational content on Twitter, business models, ownership status, and/or governance. Once the presence of a description of any one of these elements was identified, the analysis moved to the second step of coding. The second step of coding

involved breaking down and classifying each mention of a facet of an informational flow along the lines of whether it included a description of *who* information flows to, *what* information is flowing, *when* the information is flowing, *where* the information is flowing, *how* the information is flowing, or *why* the information is flowing. After this second tier of coding and categorization was complete, within each second tier code, a third tier of coding took place based on in situ analysis. In this third step, the results of the second-stage coding were analyzed to inductively discover recurring descriptions of, for example, particular actors named as "who" information flows to, or recurring discussions about "where" information flows. In essence, the first tier classifies the instance of a description of a facet of information flow, the second tier explores the descriptive details of the instance, and the third tier identifies prevalent trends across the descriptive details.

In this section, I have chosen to prioritize and organize the themes observed during coding based on the second tier of the coding scheme (who, what, when, where, how, and why) rather than through the first tier (data/metadata, algorithms, protocols, interfaces, defaults, users, types of informational content on Twitter, business models, ownership status, and/or governance). This is because, first, Twitter, Inc. does not present its discussion of the facets of information flows as neatly isolated facets. Instead, the descriptions of information flow that Twitter, Inc. provides to users frequently discuss multiple components of information flows in tandem. For example, discussions about advertisers were often co-present with descriptions of particular types of user data. Second, van Dijck's techno-cultural and socio-economic dimensions operate as broad categories, and this analysis is now moving to explore how Twitter, Inc. describes the

specific, prevalent details of how information flows on Twitter to registering users. For example, which advertisers specifically and recurrently are being discussed and which pieces of user-data are being described as flowing to advertisers. Lastly, an organizational scheme built from the second-tier of the coding scheme makes it easier to discuss the emergent trends in coherent and logical groupings. While van Dijck's framework serves as a first-order tool for identifying the presence of a description of an information flow, that is only the first step in the analysis.

The themes discussed in this section became visible from the application of the coding scheme itself. However, after the coding process, a number of additional trends in the functional presentation of the text became apparent, despite the fact that there was no initial plan to address those elements. While not formally part of the coding scheme, these additional trends also may have relevance for how users glean *principles-knowledge* from the texts. These additional trends are commented on in the *Context: Presentation of the Text* section, near the end of this chapter.

Who is involved with information flows?

Throughout the corpus, there are a number of descriptions of different senders and receivers involved in information flows on Twitter. This section will provide an overview of the emergent themes of who the texts describe as involved with information flows. In terms of this dissertation's framework of data/metadata, algorithms, protocols, interfaces, defaults, informational content, users, governance, business practices, and ownership, these themes were predominant within descriptions of the "users," "business models," and "ownership" facets. The themes are presented here organized under the headers of

"You;" "Twitter;" "Others, Registered Users, and Unregistered Users" and "Third-Parties."

You.

One of the most common explicitly or implicitly referenced senders and/or receivers of information on the Twitter platform within the corpus is "you." In every single document included in the corpus, there is at least one instance of "you." For example, the Twitter.com landing page proclaims, "Start a conversation, explore *your* interests, and be in the know [emphasis added]." This statement implies "you" as both a possible creator of informational content (starting a conversation) and as a receiver of informational content (conversation, information related to your interests, information that allows you to be in the know).

This construction of "you" also projects an image of who the other "yous" on the service (other users) might be. These other "yous" may also be interested in conversation, "your" interests, or may have the information that would put one in the know. Interestingly, in ascribing *interests* to "you," this text projects an image of Twitter users as entities capable of having interests (i.e., of being human). Twitter's Privacy Policy, in fact, explicitly refers to *people* it is first line: "Twitter instantly connects people everywhere to what's most meaningful to them." Furthering the connection between "you" and personhood, the Terms of Service states: "All Content, whether publicly posted or privately transmitted, is the sole responsibility of the person who originated such Content." These descriptions may not convey a robust picture of all Twitter users however as, in 2014, 8.5% of Twitter's user base was actually composed of "bots" or

otherwise automated accounts (Crum, 2014), something a reader would not discover from the corpus.

Across the entire corpus, there is only one robust articulation of what kind of entities act as the "you" behind registered accounts. The Terms of Service provides this important definition when it states:

> You may use the Services only if you can form a binding contract with Twitter and are not a person barred from receiving services under the laws of the United States or other applicable jurisdiction. If you are accepting these Terms and using the Services on behalf of a company, organization, government, or other legal entity, you represent and warrant that you are authorized to do so.

In this statement, the text establishes that "you," are a person, potentially representing a company, organization, government, or some other type of legal entity. However, Twitter, Inc. smoothes over some of the differences between these different "yous" in its Terms of Service when it states, "What you say on Twitter may be viewed all around the world instantly. You are what you Tweet!" This is to say, "you" is constituted by the informational content that "you" sends. The labels of "company," "organization," and "government" are washed away as each becomes reduced to the content they create. This flattening of different users to the output they produce and the lack of recurring descriptions of who might occupy the "you" position may potentially impact a reader's ability to robustly develop their *principles-knowledge* of who functions as a "you" on the site.

Twitter, Twitter, Inc., Our, and We.

One of the more interesting, but challenging aspects of the coding process related to how to treat the occurrence of the term "Twitter." The difficulty arose because across the corpus, there were multiple instances where the term "Twitter" was used to refer to a

technology, and multiple instances where the term was used to refer to the business that operates the technology (what this dissertation has referred to as "Twitter, Inc."). For example, in the Terms of Service, the first full sentence of the document reads:

> These Terms of Service ("Terms") govern your access to and use of the services, including our various websites, SMS, APIs, email notifications, applications, buttons, and widgets, (the "Services" or "Twitter"), and any information, text, graphics, photos or other materials uploaded, downloaded or appearing on the Services (collectively referred to as "Content").

Here, "Twitter" is defined as services comprised by various websites, SMS, APIs, email notifications, applications, buttons, and widgets. Yet later, the Terms of Service states:

> You agree that this license includes the right for Twitter to provide, promote, and improve the Services and to make Content submitted to or through the Services available to other companies, organizations or individuals who partner with Twitter for the syndication, broadcast, distribution or publication of such Content on other media and services, subject to our terms and conditions for such Content use.

Here, the use of "Twitter" appears to refer to the business entity that purveys the technology, an organization to which "you" can give legal rights regarding content. This dual-use of the term "Twitter" creates ambiguity and at points collapses some of the conceptual boundary between Twitter the technology and Twitter, Inc. the business entity.

The phrase "we" or "our" is also used heavily in the Terms of Service, Privacy Policy, and the Twitter Rules, to refer to Twitter, Inc. It is, however, occasionally used in a way that could confuse Twitter, Inc. for the Twitter platform. For example, the Privacy Policy states, "Our default is almost always to make the information you provide public for as long as you do not delete it from Twitter, but we generally give you settings to make the information more private if you want." The phrase "our default" is somewhat confusing as Twitter, the technology, contains default settings for Tweets, but the

statement appears to be more readily about the prerogative of Twitter, Inc. Such conflation could make it difficult for a reader to develop *principles-knowledge* about the platform, about defaults, and about the relationship between the platform's owner's prerogatives and the platform itself.

In a handful of places within the corpus, passive sentence construction additionally obscures Twitter's (or Twitter, Inc.'s) role as the sender or receiver of information. For example, in the section titled, "Tweets, Following, Lists, and other Public Information," the Privacy Policy states, "Your public information is broadly and instantly disseminated." The subject of this sentence (your public information) does not perform the action (dissemination). Instead, the passive construction of this sentence draws significance away from the agent who acts, thus partially obscuring the actor that functions as the sender within information flows.

"Others," Registered Users and Unregistered Users.

There are a number of descriptions in the corpus of whom information comes from and to whom information goes to outside of the realm of "you" and "Twitter." Occasionally, these descriptions are quite explicit, providing specific names of recipients. For example, the Privacy Policy states, "For instance, your public user profile information and public Tweets may be searchable by search engines and are immediately delivered via SMS and our APIs to a wide range of users and services, with one example being the United States Library of Congress, which archives Tweets for historical purposes." The Library of Congress is named explicitly in this statement as a receiver of Tweets, though this single spot in the middle of the privacy policy is the only place in the corpus where it is ever discussed. Simultaneously, however, this statement also describes

a vaguely bounded set of recipients, such as search engines and "a wide range of users and services." Absent is a discussion of what a "range of users" means practically. Are these users individuals? Businesses? Researchers? Government entities? Data miners? Taken as a whole this statement provides only a singular specific point of focus (the Library of Congress) against an otherwise hazy field of receivers (search engines and a wider range of users and services). This lack of specificity may inhibit a reader's ability to build up a mental model of who is practically involved as a receiver of information flows on Twitter.

Most descriptions of the recipients of information in the corpus tend towards the vague rather than the explicit. "Others" is one of the most frequently named recipient in the corpus. For example, the new-user sign-up page states at the very bottom that, "Note: *Others* will be able to find you by name, username or e-mail [emphasis added]." This language does not specify whether these "others" might include other registered users, unregistered users, search engines, advertisers, governments, etc. "Others" are also mentioned in several locations within the Privacy Policy, the ToS, and in the user interface. For example, the ToS states, "This license is you authorizing us to make your Tweets available to the rest of the world and to let *others* do the same [emphasis added]" Again, whom these "others" are that a user is authorizing to make their Tweets available is left ambiguous. The addition of greater specificity might better help newly registering users in building their *principles-knowledge.* With such *principles-knowledge,* a user may be better positioned to make a more informed decision about what information to provide, such as whether to use a real name, a pseudonym, or whether to sign-up at all.

In a few locations, "others" gains a bit more specificity when it narrows down to "other users" as the senders/recipients of informational content. For example, the Terms of Service states, "You are responsible for your use of the Services, for any Content you provide, and for any consequences thereof, including the use of your Content by *other users* and our third party partners [emphasis added]." Later in the corpus, "other users" becomes gains even greater clarity as the distinction is made between registered users and non-registered users. Although they are not called "unregistered users" (as this dissertation has called them), the corpus does distinguish between the parts of the Twitter interface and content creation protocols that registered users can access versus what non-registered users have access to. For example, the Privacy Policy states that, "Any registered user can send a Tweet, which is a message of 140 characters or less that is public by default and can include other content like photos, videos, and links to other websites" but later that, "Some Services, such as search, public user profiles and viewing lists, do not require registration." This statement does serve as a subtle indication that some information on Twitter is made accessible to those beyond the millions of registered Twitter users. However, again, the description of the access is incomplete. While search, public user profiles, and viewing lists are indicated as services that an unregistered user may access, the phrase "some services" is vague. Despite the fact that there are also different "types" of users on the site, such as "verified users," those distinctions are never discussed within the corpus.

Third-parties, advertisers, and "our partners."

The last prominent theme in descriptions of senders and receivers of information within the corpus deals with "third-parties," "advertisers," and "our partners." These

terms appear most prominently in the policy documents. For example, under the "Information Collection and Use" section of the Privacy Policy, there is an entire section devoted to "Third-Parties." This section begins:

> Twitter uses a variety of third-party services to help provide our Services, such as hosting our various blogs and wikis, and to help us understand the use of our Services, such as Google Analytics. These third-party service providers may collect information sent by your browser as part of a web page request, such as cookies or your IP address. Third-party ad partners may share information with us, like a browser cookie ID or cryptographic hash of a common account identifier (such as an email address), to help us measure ad quality and tailor ads.

This statement is fairly typical of the way that third-parties, and the kinds of information they receive/send, is described throughout the corpus. While this statement does contain one specific example of a third-party that Twitter partners with (Google Analytics), and specific types of information that a third-party might receive (browser cookie ID, e-mail address), it is vague about the exact limits of the information collected and shared, and about what these third-parties might do with the information. For example, the Terms of Service state, "You are responsible for your use of the Services, for any Content you provide, and for any consequences thereof, including the use of your Content by other users and *our third party partners*." Yet, not a single third-party partner outside of Google Analytics is ever named in the statement, so connecting "consequences" to the actions these parties might take would seem incredibly difficult. As a result, the vagueness of the description may impede users in their ability to develop robust *principles-knowledge* about who third-parties are, how these third-parties use content from Twitter, and what the "consequences" of such use might look like.

The fact that Twitter, Inc.'s information exchange with ad-partners and third-party services has a commercial dimension is also not explicitly discussed. While the corpus

does address the fact that Twitter has advertisements, and works with advertisers and third-party partners, it is never explicitly stated how this is tied to Twitter, Inc.'s business models. The corpus does not describe, for example, Twitter's partners in Twitter's "Certified Products" program, partners such as "SocialBro" which, "gives business powerful tools to analyze and act on Twitter data to grow, engage and monetize their audience" (twitter.com, 2014j) or "DataSift" which, "provides both realtime and historical access to the full Twitter firehose" and " … offers a flexible pricing scale that makes enterprise-level data accessible to companies of any size" (twitter.com, 2014c). Instead, these aspects of the information flows on Twitter are left unstated in the corpus. Thus, users may be put at a disadvantage in building an understanding of the third-party data vendor ecosystem that exists which capitalizes on users' content creation.

What information flows?

Across the corpus, there are dozens of descriptions of what information flows within the Twitter platform. This section will provide an overview of the themes and trends that emerged after the application of the coding scheme. In terms of this dissertation's information flow framework of data/metadata, algorithms, protocols, interfaces, defaults, informational content, users, governance, business practices, and ownership, these themes were predominant within descriptions of the "informational content," "data/metadata," and "governance" facets. Trends that emerged within the analysis are grouped together in this section under the headers of "informational content," "optional, personal, and public information," and "metadata."

Informational Content.

Many of the descriptions of information within the corpus focus on specific kinds of content that Twitter enables a user to access. Most prominently, Twitter's landing page suggests that a user can "start a conversation" or "explore your interests." Both of these descriptions imply a specific kind of informational content one might find through the platform. This is a very different framing of what one finds on Twitter than a statement such as "start Tweeting today" or "explore Tweets." The former descriptions essentially privilege message over medium rather than the other way around.

The Twitter Rules focus almost exclusively the kinds of informational content allowed and disallowed within the platform. The document begins by stating that with limited exception, "we do not actively monitor and will not censor user content, except in limited circumstances described below." Forbidden from the services are forms of content such as impersonation, materials that are others' intellectual property, direct threats of violence, abuse and spam, and pornographic profile photos or header photos (to name a few). In a few cases, the document provides examples of what might constitute content forbidden from the service. For example, the Twitter Rules state, "You may not publish or post other people's private and confidential information, such as credit card numbers, street address or Social Security/National Identity numbers, without their express authorization and permission." Here, the private and confidential information of others is defined as a form of content forbidden from the service and a handful of examples are provided. However, while credit card numbers and social security numbers are perhaps fairly commonly treated as private and confidential, street address may not be. In fact, street address is often considered public information. This listing calls into question what other types of informational content might qualify as the private and

confidential information of others. For example, would someone else's height, weight, marital status, or publicly accessible court records qualify? While this document provides a general sketch of the governance of informational content, at the same time it also contains a somewhat hazy articulation of the boundaries of these forms of content. Further, it does not provide an articulation of how procedurally Twitter, Inc. makes a determination about whether a given piece of information falls under a specific category of banned informational content. As a result, users may face hurdles in developing this component of *principles-knowledge* regarding the governance of informational flow on the platform.

Optional, public, and private information.

The Privacy Policy provides a detailed gradation of different kinds of information *about users* within certain information flows on Twitter. In the section, "Information Collection and Use," the Privacy Policy outlines nine general categories of information: "Information Collected Upon Registration," "Additional Information," "Tweets, Following, Lists and other Public Information," "Location Information," "Links," "Cookies," "Log Data," "Widget Data," and "Third-Parties." In its discussion of these nine categories of information collection and use, the Privacy Policy uses the terms "optional," "personal," and "public" to describe groups of information within each category. In its section on "Information Collected Upon Registration," for example, the Privacy Policy states:

> When you create or reconfigure a Twitter account, you provide some *personal information*, such as your name, username, password, and email address. Some of this information, for example, your name and username, is listed *publicly* on our Services, including on your profile page and in search results [emphasis added].

From this, a reader may learn that name, username, password, and email address are considered to be "personal information." While name and username are given the status of "public" information in addition to "personal," the status of passwords and e-mail addresses are left unaddressed (although the statement *implies* that they are non-public). In terms of the optional nature of this information, Twitter states, "When you create or reconfigure a Twitter account, *you provide*… [emphasis added]." While this statement does not explicitly use the word "non-optional" or "mandatory," there is no hedging about whether or not a user provides it. In other sections, such as the "Additional Information" section, the Privacy Policy lists a number of pieces of information that are optionally submitted by a user by stating: "You *may* provide us with profile information to make public, such as a short biography, your location, your website, or a picture [emphasis added]." The phrase "may" is the distinguishing characteristic in the sentence that points to the optional nature of the submission.

The nine different sections of the Privacy Policy identify at least 43 different types of information about users that can exist within Twitter. However, in many cases, whether this information is optional or public is not stated clearly, or it is left implied. Table 69 provides a charting of the individual pieces of information identified, the section it is discussed in, and whether it is described as optional or public, explicitly or implicitly.

Table 69

Listing of Informational Content Discussed in Privacy Policy and Optional/Public Status

Information	**Section of Privacy Policy**	**Optional?**	**Public**
Name	Information Collected Upon Registration	No - explicit (no use of "may")	Yes - explicit
Username	*supra*.	No - explicit (no use of "may")	Yes - explicit
Password	*supra*.	No - explicit (no use of "may")	No - implicit
E-Mail Address	*supra*.	No - explicit (no use of "may")	No - implicit
Short Biography (Profile)	Additional Information	Yes - explicitly (use of "may")	Yes - explicit
Location (Profile)	*supra*.	Yes - explicitly (use of "may")	Yes - explicit
Your Website (Profile)	*supra*.	Yes - explicitly (use of "may")	Yes - explicit
Picture (Profile)	*supra*.	Yes - explicitly (use of "may")	Yes - explicit
Cell Phone Number	*supra*.	Yes - explicitly (use of "may")	Not stated explicitly or implicitly
Your Address Book	*supra*.	Yes - explicitly (use of "may")	Not stated explicitly or implicitly
Registration Info on other services	*supra*.	Yes - explicitly (use of "may")	Not stated explicitly or implicitly
Profile Information from other services	*supra*.	Yes - explicitly (use of "may")	Not stated explicitly or implicitly
Tweets	Tweets, Following, Lists and other Public Information	Yes - Implicit	"Our default is almost always to make the information you provide public for as long as you do not delete it from Twitter, but we generally give you settings to make the information more private if you want"
Metadata provided with Tweets, such as when you tweeted	*supra*.	Not stated explicitly or implicitly	*supra*.
Lists you create	*supra*.	Yes - Implicit	*supra*.
People You Follow	*supra*.	Yes - Implicit	*supra*.

Information	**Section of Privacy Policy**	**Optional?**	**Public**
Tweets You Mark as Favorite	*supra*.	Yes - Implicit	*supra*.
Tweets You Retweet	*supra*.	Yes - Implicit	*supra*.
"Many Other Bits of Information That Results from Your use of the Services"	*supra*.	Not stated explicitly or implicitly	*supra*.
Photos You Submit	*supra*.	Yes - Implicit	*supra*.
Videos You Submit	*supra*.	Yes - Implicit	*supra*.
Links You Submit	*supra*.	Yes - Implicit	*supra*.
Location in your Tweets	Location Information	Yes - explicitly	Not stated explicitly or implicitly
Location in your Profile	*supra*.	Yes - explicitly	Not stated explicitly or implicitly
Trend Location	*supra*.	Yes - explicitly	Not stated explicitly or implicitly
How you interact with links across our service	Links	Not stated explicitly or implicitly	Not stated explicitly or implicitly
Session Cookie	Cookies	Yes, but some Services may not function properly if you disable cookies	Not stated explicitly or implicitly
Persistent Cookie	*supra*.	Yes, but some Services may not function properly if you disable cookies	Not stated explicitly or implicitly
IP address	Log Data	No - implied	Not stated explicitly or implicitly
Browser Type	*supra*.	No - implied	Not stated explicitly or implicitly
Operating System	*supra*.	No - implied	Not stated explicitly or implicitly
Referring Web Page	*supra*.	No - implied	Not stated explicitly or implicitly
Pages Visited	*supra*.	No - implied	Not stated explicitly or implicitly
Location	*supra*.	No - implied	Not stated explicitly or implicitly
Your Mobile Carrier	*supra*.	No - implied	Not stated explicitly or implicitly

Information	**Section of Privacy Policy**	**Optional?**	**Public**
Device	*supra*.	No - implied	Not stated explicitly or implicitly
Application IDs	*supra*.	No - implied	Not stated explicitly or implicitly
Search Terms	*supra*.	No - implied	Not stated explicitly or implicitly
Cookie Information	*supra*.	No - implied	Not stated explicitly or implicitly
Web page you visited	Widget Data	Yes - implied	Not stated explicitly or implicitly
Cookie that identifies your browser	*supra*.	Yes - implied	Not stated explicitly or implicitly
Browser cookie ID	Third-Party	Yes - implied	Not stated explicitly or implicitly
Cryptographic hash of common account identifier	*supra*.	Yes - implied	Not stated explicitly or implicitly

The implicit nature of many of the descriptions may inhibit the development of *principles-knowledge* regarding whether Twitter, Inc. treats that information as optional or public.

Following the section on information collection and use comes a section on "Information Sharing and Disclosure." This section begins with the "Tip" that, "We do not disclose your private personal information except in the limited circumstances described here." However, the word "private" does not appear in the entire previous "Information Collection and Use" section. In fact, within the entire "Information Sharing and Disclosure" section, there is no explicit definition of what information is considered "private personal information" and what information is not. Instead, a reader is left to interpret what constitutes this category of information based on a sub-section entitled, "Non-Private or Non-Personal Information." This section states:

> We may share or disclose your non-private, aggregated or otherwise non-personal information, such as your public user profile information, public Tweets, the people you follow or that follow you, or the number of users who clicked on a particular link (even if only one did).

A reader is left to construe that if it is not listed as non-private or non-personal, then it must conversely be private and personal. However, the use of "such as" suggests that this is not the only information considered to be non-private or non-personal. Thus, the open-ended nature of the statement makes it difficult to interpret what exactly counts as private, personal information.

The phrase "private information" does appear elsewhere in the corpus. The Twitter Rules state in their discussion on prohibited content, "You may not publish or post other people's *private and confidential information*, such as credit card numbers, street address or Social Security/National Identity numbers, without their express authorization and permission [emphasis added]." However, as previously noted, while credit card numbers and social security numbers are frequently recognized as forms of private information, street addresses are not as widely recognized as such. This also raises an interesting question about whether location information included in Tweets could be considered "private" or "confidential" information. For example, if a friend came over to your house and posted a Tweet from your driveway that said, "About to grab some coffee with @yourusername, picking them up from their driveway right now!" and included the geo-location coordinates (as allowed by Twitter's web-interface), would that constitute someone's private or confidential information? The ambiguity about what constitutes private information versus public information may function as a hurdle in individuals' development of *principles-knowledge* regarding how certain pieces of information are treated within the platform.

Metadata.

A small number of locations within the corpus address metadata. Generally, discussions of metadata are subsumed within discussions of informational content. For example, the Terms of Service lays out a definition of "Content" that explicitly defines informational content while also implicitly including metadata as Content:

> These Terms of Service ("Terms") govern your access to and use of the services, including our various websites, SMS, APIs, email notifications, applications, buttons, and widgets, (the "Services" or "Twitter"), and any information, text, graphics, photos or other materials uploaded, downloaded or appearing on the Services (collectively referred to as "Content")."

This definition collapses the distinction between the kinds of content that a user might upload consciously (such as the 140 characters of a Tweet or a picture attached to a Tweet) together with the metadata surrounding the informational content that appears on the services, despite the fact that a user may be unaware this information is automatically generated.

In fact, metadata is only explicitly mentioned once within the corpus. In the section entitled, "Tweets, Following, Lists, and other Information" the Privacy Policy states:

> Our Services are primarily designed to help you share information with the world. Most of the information you provide us is information you are asking us to make public. This includes not only the messages you Tweet and the metadata provided with Tweets, such as when you Tweeted, but also the lists you create, the people you follow, the Tweets you mark as favorites or Retweet, and many other bits of information that result from your use of the Services.

Further, this statement only describes one kind of metadata (time of Tweet), despite the fact that, as Chapter 4 illustrates, there are more than 30 fields that constitute a "Tweet." The majority of the different types of metadata that constitute a Tweet instead go unaddressed in the corpus.

When does information flow?

On the whole, descriptions of the temporal aspects of information flow were not as common as descriptions of what information might be flowing or who is involved with information flows. Despite this, there were three noteworthy trends that emerged from the coding process. When descriptions of the temporal nature of information flows were present, they generally focused on: descriptions of the immediate or real-time nature of information flows on the platform, descriptions of what users can change about information flows at a future point, and descriptions of the length of time Twitter stores and keeps certain kinds of content. In terms of this dissertation's information flow framework of data/metadata, algorithms, protocols, interfaces, defaults, informational content, users, governance, business practices, and ownership, the descriptions primarily came from the facets of "informational content," "interfaces," and "defaults."

Right now.

A majority of the descriptions of time in relation to information flows focus on the "instant," "immediate," or "real-time" nature of the platform itself or the informational content available through the platform. The policy documents tend towards describing the temporal nature of the affordances of the platform, whereas the landing page and interface focus more on the timeliness of the informational content. For example, the Privacy Policy starts off by stating, "Twitter *instantly* connects people everywhere to what's most meaningful to them [emphasis added]." This prioritizes the expedience at which the platform delivers information over the timeliness of the informational content. The landing page on the other hand, entices individuals to sign-up with Twitter with the promise of access to real-time *information* rather than just a real-

time *platform*. For example, Twitter's landing page implies the contemporary nature of the informational content within Twitter when it states, "Start a conversation, explore your interests, *be in the know* [emphasis added]." This statement implies a level of temporal exclusivity to the content on the service.

The Twitter.com interface further reinforces the "fresh" nature of the informational content. In addition to structuring the timeline inside the interface with the newest Tweets on top, the "Discover" tab of the user interface declares it contains "What's happening *right now*, tailored for you [emphasis added]." The promise of information flows even further mired in the real-time is used as an appeal for users to provide even more information about themselves. As part of the mobile settings page of the Twitter.com interface, an appeal for users to include their mobile phone number states: "Expand your experience, get closer, and stay *current*." Taken together, these statements reinforce the concept of the immediacy of the Twitter platform and the contemporary nature of the informational content.

Later or at any time (but right now).

Across the corpus, there are a handful of instances where the documents deal with future events. In most cases, these statements occur when users are informed about how they can either delete informational content or when they are told that they can change the default settings either "later" or "at any time." For example, the Privacy Policy states that a user may share their e-mail contact list with Twitter in order to find people they may know, and that, "We may *later* suggest people to follow on Twitter based on your imported address book contacts, which you can delete from Twitter *at any time* [emphasis added]." Similarly, on the new user-registration page, the small light grey text underneath

the "Create my account" button states, "Note, others will be able to find you by name, username or email. Your email will not be publicly shown. You can change your privacy settings *at any time* [emphasis added]."

While the new user registration page suggests an individual can change their privacy settings at any time, that statement is not entirely accurate. Users are unable to select their privacy settings or change information flow defaults (such as choosing to have a protected account) during the sign-up process itself. Instead, they must wait until after the registration and orientation process and must then figure out how to navigate to the settings page in order to change the defaults. As the ability to access the settings pages (or to even know what the settings pages looks like) is dependent on having a registered account, individuals contemplating adoption of the technology are put at a disadvantage in their ability to develop *principles-knowledge* about these elements until after registration.

Retention and information lifecycles.

Within the corpus, there are a handful of descriptions of the temporality of information flows that address when information is retained, stored, and/or removed from Twitter. Although most descriptions of information content within the corpus focus on its active transmission, a handful of mentions focus on retention using terms such as "storage" or "store" or "archive." For example, the Privacy Policy states, "You understand that through your use of the Services you consent to the collection and use (as set forth in the Privacy Policy) of this information, including the transfer of this information to the United States and/or other countries for *storage*, processing and use by Twitter."

While mentions of storage are few, an even smaller number lay out explicit timelines for retention. For example, an explicit lifecycle is never given for Tweets. One is instead implied by the statement in the Privacy Policy, reading: "Our default is almost always to make the information you provide public *for as long as you do not delete it* from Twitter." From this, one might gather that Twitter stores Tweets indefinitely. However, the ToS appears to throw this interpretation into some question when it states, "… Twitter may stop (permanently or temporarily) providing the Services (or any features within the Services) to you or to users generally and may not be able to provide you with prior notice. We also retain the right to create limits on use and storage at our sole discretion at any time without prior notice to you." This wording suggests that Twitter, Inc. retains the right to limit how much content it maintains or for how long it maintains that content, and that Twitter, Inc. may not give notice to users if it does change its retention practices. As a result, the *principles-knowledge* these descriptions foster seems to entail a measure of uncertainty.

While Twitter may keep user content public until the user deletes it, deleting Tweets is sometimes no easy task. Twitter does not offer a way to bulk delete Tweets. Instead, if a user wants to delete Tweets, they must do so from their timeline, one at a time. Even if a user deletes their entire account, Twitter still maintains the Tweets that user generated for some time. The Privacy Policy states:

> You can also permanently delete your Twitter account. If you follow the instructions here, your account will be deactivated and then deleted. When your account is deactivated, it is not viewable on Twitter.com. For up to 30 days after deactivation it is still possible to restore your account if it was accidentally or wrongfully deactivated. After 30 days, we begin the process of deleting your account from our systems, which can take up to a week.

As a result, if a user deletes their account, their Tweets may still stay available for some period of time. Although the corpus does not address this, if other users have quoted portions of a Tweet or have retweeted a user by copying and pasting their Tweets, that information will still stay in Twitter's system.

Log data and widget data are examples of data described as having explicit lifecycles in the corpus. However, even these timelines are hedged and given caveats. In its section on "Widget Data," the Privacy Policy, states:

> We may tailor content for you based on your visits to third-party websites that integrate Twitter buttons or widgets. When these websites first load our buttons or widgets for display, we receive Log Data, including the web page you visited and a cookie that identifies your browser ("Widget Data"). After a *maximum of 10 days*, we start the process of deleting or aggregating Widget Data, *which is usually instantaneous but in some cases may take up to a week*. While we have the Widget Data, we may use it to tailor content for you, such as suggestions for people to follow on Twitter. [emphasis added]

Similarly, the section in the Privacy Policy on "Log Data" states, "…we will either delete Log Data *or* remove any common account identifiers, such as your username, full IP address, or email address, after 18 months [emphasis added]." While these statements do provide a general timeline, they are built with some level of flexibility. The first statement lays out a range of time in which deletion may occur. The second statement suggests that log data is either deleted or certain aspects of the log data are anonymized, but a user has no idea of knowing which has actually taken place. These descriptions of information lifecycles lack specificity and precision and as a result, a user may never be entirely certain of when exactly information removal has actually taken place.

Where information flows.

Despite the fact that "location" was one of the least discussed elements of information flows within the corpus, there were a handful of trends that emerged from the

few descriptions present. First, descriptions of "where" information flows are frequently entwined with descriptions of "who" is involved with information flows throughout the corpus. For example, in multiple locations, "The World" is described as the recipient of information while simultaneously a location where information is sent. Second, when the legal regulation of informational content on Twitter is discussed, specific jurisdictions such as the U.S. are repeatedly mentioned as a location where information flows. Next, other websites come up frequently as locations where Twitter content and aspects of the Twitter interface itself may appear. Finally, Twitter itself was sometimes described as a kind of environment or place within the corpus. In terms of this dissertation's information flow framework of data/metadata, algorithms, protocols, interfaces, defaults, informational content, users, governance, business practices, and ownership, these trends primarily come from facets of "data/metadata," "users," "governance," and "interfaces."

The world.

Perhaps the most grandiose description of location in relation to information flows on Twitter occurs in discussion about "The World." This appears both implicitly and explicitly throughout the corpus. For example, it appears implicitly on the Twitter.com landing page. There, to the right of the relatively simple text, "Welcome to Twitter. Start a conversation, explore your interests, and be in the know" is an image of an iPhone running the Twitter app. Open on this app is a Tweet containing a picture of Austrian skydiver Felix Baumgartner, who is getting ready to jump from the Red Bull Stratos helium balloon, 24 miles above earth.[30] While Baumgartner is the subject of the photo, the planet Earth comprises a clearly visible background. In choosing this particular Tweet and this particular image, Twitter, Inc. is invoking a visual connection between

[30] https://twitter.com/redbullstratos/status/257986797324345344

information flows from mobile (via the iPhone), brands (in choosing a Tweet from Red Bull), communication at a distance (the photograph of a man, hovering above earth, yet appearing on Twitter), and the ability to communicate with the world itself (through the image of the earth).

Rather than implicitly, both the Privacy Policies and the ToS directly state the "world" as a recipient of information. The ToS states, "What you say on Twitter may be viewed *all around the world* instantly. You are what you Tweet! [emphasis added]" and, in a section on users' rights, states "This license is you authorizing us to make your Tweets available *to the rest of the world* and to let others do the same [emphasis added]" (again, here the use of "others" leaves the exact actors ambiguous). The Privacy Policy goes so far as to say, "Our Services are primarily designed to help you share information with the world." While there is no denying that Twitter has been a technology used in a number of different events across the world, is the entire world truly a recipient of information on Twitter? The service is (at the time of writing) blocked in China, Iran, and North Korea. It has also previously been blocked in Egypt and Turkey. This is, of course, without even beginning to consider the millions without connection to the technologies necessary to access Twitter, or to those who do, but do not use the service.

The United States and other jurisdictions.

The Privacy Policy, the Terms of Service, and the Twitter Rules all make mention of the fact that information flows on Twitter involve the transmission of content to many different countries. The United States is the most prominently mentioned country within the corpus (perhaps not surprisingly, given that this is where Twitter was founded and is headquartered). Specifically, there are multiple mentions of the fact that, regardless of

where the information is submitted from, it is transferred to the United States. In the Terms of Service, this is stated as:

> Any information that you provide to Twitter is subject to our Privacy Policy, which governs our collection and use of your information. You understand that through your use of the Services you consent to the collection and use (as set forth in the Privacy Policy) of this information, including the transfer of this *information to the United States and/or other countries* for storage, processing and use by Twitter.

The Privacy Policy reminds a reader of this again when it states: "Irrespective of *which country you reside in or supply information from,* you authorize Twitter to use your information *in the United States and any other country where Twitter operates* [emphasis added]."

While Twitter itself is sometimes described as a location (discussed later in this section), it is not a place where users are unmoored and untouchable by the governance of regional law. Outside of the focus on the United States, regional or local laws are mentioned a few times in the corpus. Specifically, these mentions usually function as a reminder to users that they are still bound to local laws governing the production of informational content. The Twitter Rules state, "International users agree to comply with all *local laws* regarding online conduct and acceptable content [emphasis added]." Content that breaks U.S. laws are also verboten. While the corpus does address the applicability of local laws to users' informational content generation on Twitter, the corpus does not *explicitly* address the fact that Twitter censors content within geographic regions if they are given a valid legal request to do so. For example, Twitter has agreed to block certain Tweets generated by a newspaper in Turkey from being accessed by Twitter users within Turkey (Fiveash, 2015). While Twitter does inform a user if they try to access blocked content that the content has been blocked in their region, the fact that this

can occur is not addressed within the corpus. As a result of the way that jurisdictionally based restrictions on information flows are described, it is questionable whether or not a user could develop accurate *principles-knowledge* of the fact Twitter augments flows in this capacity from the corpus.

Third-party websites.

Information uploaded to Twitter is often displayed on other websites besides Twitter.com. This is according to the Terms of Service, which states, "The Content you submit, post, or display will be able to be viewed by other users of the Services and *through third party services and websites* [emphasis added]." Widgets and buttons are the primary way this occurs. These embeddable elements of Twitter's interface facilitate the display of content from Twitter inside third-party websites.

User visits to third-party websites that have embedded widgets and buttons also generate data about the visit, which is then transmitted back to Twitter. As Twitter's Privacy Policy states:

> We may tailor content for you based on your visits to third-party websites that integrate Twitter buttons or widgets. When these websites first load our buttons or widgets for display, we receive Log Data, including *the web page you visited* and a cookie that identifies your browser ("Widget Data"). After a maximum of 10 days, we start the process of deleting or aggregating Widget Data, which is usually instantaneous but in some cases may take up to a week. While we have the Widget Data, we may use it to tailor content for you, such as suggestions for people to follow on Twitter. [emphasis added]

This is one of the only instances in the corpus where the fact that Twitter receives information about user web-browsing habits is discussed.

However, nowhere in the corpus are readers given an example or definition of what a widget or button is or looks like. While individuals with a higher level of digital literacy might be able to recognize what a widget or button is, those without this prior

knowledge may be lost. As a result, certain individuals are likely put at a disadvantage in their ability to build-up *principles-knowledge* of how information, and possibly quite sensitive information about browsing habits, flows between the third-party website locations and Twitter.

Twitter.

In a number of locations in the corpus, Twitter is described as a kind of environment or location onto itself. For example, the Terms of Service describes Twitter as a kind of naturally evolving environment, stating, "Twitter has an evolving set of rules for how *ecosystem* partners can interact with your Content. These rules exist to enable an *open ecosystem* with your rights in mind [emphasis added]." Although it is not part of this corpus, Twitter's API Guide further expands this metaphor, using the homage to the *Audubon Field Guide* to suggest Twitter as a habitat where data lives. While there are only a handful of locations in the corpus where this positioning of Twitter as a place or ecosystem occurs, it is a powerful metaphor for conceptualizing the service as a place. It is difficult to say exactly how this might impact users' relative informational power, other than drawing into question how this sense of place conflicts with the idea of Twitter as beholden to specific national jurisdictions and governance.

Why does information flow?

Across the corpus, there are a small number instances where justifications are given for why information flows occur in the arrangements they do. From these descriptions, two themes emerged. First, the corpus provides a number of justifications for why information flows in certain arrangements based on Twitter, Inc.'s "goals." Second, in a number of locations, descriptions of why information flows focus more

narrowly on their relation to "improving the services." In terms of this dissertation's information flow framework of data/metadata, algorithms, protocols, interfaces, defaults, informational content, users, governance, business practices, and ownership, the themes touched on facets of "data/metadata," "informational content," "users," and "business practices."

Twitter, Inc.'s goals.

The Twitter Rules state: "*Our goal* is to provide a service that allows you to discover and receive content from sources that interest you as well as to share your content with others [emphasis added]." Facilitating individual discovery, access to informational content from sources of interest, and allowing users to share informational content with others appear as Twitter, Inc.'s goal in providing its service. These goals, incidentally, seem to align with the statement on the Twitter.com landing page, "Start a conversation, explore your interests, be in the know." Perhaps most simply, the Privacy Policy states, "Our Services are primarily designed to help you share information with the world."

Interestingly, Twitter's goal of generating revenue is never explicitly mentioned as a reason for Twitter's existence or as a justification for the arrangement of information flows on Twitter. The fact that Twitter, Inc. generates revenue via the sale of access to the firehose of user-generated content is not addressed at all in the corpus, and the fact that advertising such as promoted Tweets, promoted trends, and promoted accounts involve a commercial transaction is left *implied.* As a result, the interconnection between users' information content production and Twitter, Inc.'s revenue generation is obfuscated from the immediate view of a reader. Thus, the descriptions of Twitter's goals

foster a kind of cognitive gap between information flows and their commodification, which may impede the development of users' *principles-knowledge.*

Better, faster, safer, and more tailored.

Improving the services, the user-experience, Twitter, Inc.'s knowledge of how its services are used, and ad delivery are some of the justifications given in the corpus for why Twitter, Inc. collects certain user information. The Cookie Use statement, for example, states: "Twitter uses cookies and other similar technologies, such as pixels or local storage, to help provide you with a better, faster, and safer experience." While the document explains what cookies, pixels, and local storage are to an extent, it does not provide further elaboration of what "other technologies" Twitter, Inc. uses. However, the document goes into greater detail about why Twitter uses cookies, pixels, local storage, and other similar technologies, remarking that justifications for the flows generally fall into one of seven categories:

- "To log you into Twitter…"
- "To protect your security…"
- "To help us detect and fight spam, abuse, and other activities that violate the Twitter Rules…"
- "To remember information about your browser and your preferences…"
- "To help us improve and understand how people use our services, including Twitter buttons and widgets, and Twitter Ads…"
- "To customize our services with more relevant content, like tailored trends, stories, ads, and suggestions for people to follow…"
- "To help us deliver ads, measure their performance, and make them more relevant to you based on criteria like your activity on Twitter and visits to our ad partners' websites…"

Perhaps not surprisingly, the corpus does not address any of the potential downsides or drawbacks for user privacy that might result from such information collection. Further, absent from these justifications is any discussion of how user tracking technologies fit into the broader picture of Twitter, Inc.'s revenue generation and business models.

Tracking is instead couched as making advertising more relevant to the user, rather than about the revenue potentially gained by Twitter, Inc. through its ability to sell refined ad targeting.

As part of the Privacy Policy, the document describes how when users click on links within Twitter, Twitter, Inc. may track this information by redirecting link clicks through a pass-through hyperlink. Twitter, Inc. justifies this practice in the Privacy Policy, stating:

> Twitter may keep track of how you interact with links across our Services, including our email notifications, third-party services, and client applications, by redirecting clicks or through other means. We do this to help improve our Services, to provide more relevant advertising, and to be able to share aggregate click statistics such as how many times a particular link was clicked on.

Again, improving the Services, the relevance of advertising, and the ability to share more detailed information is given as a reason for the practice. Interestingly though, exactly *who* Twitter, Inc. shares aggregate clicks statistics with is not mentioned in this statement, thus avoiding a description of an important aspect of this information flow.

How does information flow?

Within the corpus, there were a handful of descriptions given relating to how information flows. In terms of this dissertation's information flow framework of data/metadata, algorithms, protocols, interfaces, defaults, informational content, users, governance, business practices, and ownership, these themes generally stemmed from the "algorithms," "protocols," "interfaces," "users," "business models," and "ownership" facets. The two themes that emerged after coding are presented in this section under the headers of "Unelaborated Transmission Methods" and "Sharing."

Unelaborated transmission methods.

While the corpus addresses the fact that Twitter shares information with its partners, with advertisers, and with others, the details of *how* information is shared are frequently vague or left unexplained. For example, the first mention of the APIs appears in the Terms of Service, yet they are mentioned with no explanation of what they are, or what "API" even stands for. The ToS states, "These Terms of Service ("Terms") govern your access to and use of the services, including our various websites, SMS, APIs, email notifications, applications, buttons, and widgets…" A reader who does not already know what an API is must guess that an API is some way of interacting with the Twitter platform, based on its inclusion in a list of other ways of interacting with the Twitter platform. Later, the ToS indicates:

> Except as permitted through the Services, these Terms, or the terms provided on dev.twitter.com, you have to use the Twitter API if you want to reproduce, modify, create derivative works, distribute, sell, transfer, publicly display, publicly perform, transmit, or otherwise use the Content or Services.

Underneath this statement, the document offers a tip: "We encourage and permit broad re-use of Content. The Twitter API exists to enable this." There is no discussion of who might use the APIs or might be interested in the re-use of Content, what "API" stands for, why someone might use the API, or the fact that Twitter may generate revenue through the sale of access to Tweets via the "firehose" API.

Similarly, many of Twitter's protocols and algorithms go entirely unexplained within the corpus. For example, while tweeting and following are explained as part of the new user orientation, how the trend algorithms function is not. While an @reply button and a "favorite" button appears underneath of Tweets in the Timeline as part of the user-interface, there is no explanation given of how @replies do or do not appear within the timelines of others depending on the following/follower relationship, nor what the

purpose or function of "favorites" are. The new-user orientation process appears to be oriented towards explaining tweeting, getting users connected to a few popular "accounts," people they know, and getting users to add some basic profile information. The process appears much more focused on the development of basic *how-to knowledge* around content creation (how to tweet, how to follow) rather than more in-depth *principles-knowledge* of how information flows through the platform. It bears mentioning that Twitter does offer a "Help Center," which provides some level of further detail about these subjects, but it is not included as part of the corpus as it is not a part of the new user registration process.

Sharing.

Across the corpus, there are multiple instances where information flows are described as "sharing." Twitter is designed for users to "share information with the world" as the Privacy Policy states. However, Twitter, Inc.'s commercial information exchange practices are also described as "sharing." For example, the Privacy Policy indicates: "We may share your private personal information with such service providers subject to confidentiality obligations consistent with this Privacy Policy, and on the condition that the third parties use your private personal data only on our behalf and pursuant to our instructions" and "Third-party ad partners may share information with us, like a browser cookie ID or cryptographic hash of a common account identifier (such as an email address), to help us measure ad quality and tailor ads." These information flows are never described using terms such as "bought," "sold," or "rented." However, the Privacy Policy does state, "In the event that Twitter is involved in a bankruptcy, merger, acquisition, reorganization or sale of assets, your information may be sold or transferred

as part of that transaction." This statement projects sale of "your information" as something that is a possibility only in the future. "Your information" appears to be only "shared" for now. Thus, these descriptions of how information flows on Twitter may foster a cognitive gap between Twitter, Inc.'s information exchange and its business models.

Context: Presentation of the Text

While not falling explicitly into the textual analysis above, a handful of additional trends stood out during and after the coding process. These elements can be considered as contextual factors in the presentation of the text that may additionally impact users' abilities to develop *principles-knowledge* about the information flows on Twitter. These three trends are described in this section as: hyperlinks and information accessibility in the corpus, accessibility of account settings during the sign-up process, and finally, variations of text size and color within the corpus.

Hyperlinks and information accessibility.

Twitter uses hyperlinks extensively throughout the "Join Twitter Today" page, its ToS, Privacy Policy, the "Twitter Rules," and the Cookie Use Policy documents. At times, these hyperlinks appear to put more steps between a user and descriptions of certain aspects of information flows on the platform. For example, on the "Join Twitter Today" page, after a user provides a full name, e-mail address, password, username, and makes a selection about whether they want to stay logged in on this machine, and whether they want Twitter tailored based on their recent website visits, a statement appears that, "By clicking the button [which reads 'Create my account'], you agree to the terms below." Only the first two lines of the ToS appear below this. A reader must

actively click on the textbox to expand the ToS to a larger size in order to read it. While Twitter still makes the ToS available within the page, it does add the additional burden of making the reader click *into* it. A collapsing text box such as the one that contains the ToS does not appear anywhere else in the corpus (and, from what I have come across, nowhere else on Twitter.com).

Further, the Terms of Service state:

> These Terms, the Twitter Rules and our Privacy Policy are the entire and exclusive agreement between Twitter and you regarding the Services (excluding any services for which you have a separate agreement with Twitter that is explicitly in addition or in place of these Terms), and these Terms supersede and replace any prior agreements between Twitter and you regarding the Services.

Yet, the text of the Twitter Rules and Privacy Policy do not actually appear within the ToS. As a result, while an individual signing up "agrees to the terms below," the text contained in the textbox are not the *complete* terms of the agreement. To put together the complete agreement, an individual signing up would additionally need to click on the printable version of the Privacy Policy, and click from either the Terms of Service or Privacy Policy to the Twitter Rules, which is not linked with the other "printable" versions of the policy documents. As a result, this hyperlinking puts several steps between a reader and the full terms of the agreement; a design layout that may inhibit the individual's ability to easily develop *principles-knowledge* from the policy documents.

Accessibility of settings.

As discussed in Chapter 5, the documents that appear in this corpus were selected because they contain the language that a user would encounter signing up for Twitter. Yet, within the Terms of Service and Privacy Policies—documents a user would encounter before they officially had an account—there are multiple instances of

hyperlinks to locations that one can only access *after* they have a registered account. For example, the Privacy Policy states in its section on "Modifying Your Personal Information:" "If you are a registered user of our Services, we provide you with tools and account settings to access or modify the personal information you provided to us and associated with your account." Within the sentence, the phrase "account settings" is hyperlinked to the URL: https://twitter.com/account/settings. However, an individual who is still in the sign-up process cannot visit this page, as they must be logged in with an account in order to view it. As a result, an individual who might be contemplating Twitter's account settings or trying to learn about the settings *before* signing up is going to be put at a disadvantage in being able to find out what those account controls are. Further, despite the fact that newly registering users can provide information about themselves to Twitter as part of the new-user orientation process, they do not have the ability to set privacy and security controls *before* they create their account or *as* they create the account.

Text size and color.

In a number of locations within the corpus, the documents contain descriptions of information flows in smaller font sizes and lighter colors than the rest of the text. As a result, these descriptions may not be as readily apparent to users (particularly users with certain kinds of vision problems). For example, on the new user sign-up page, despite the fact that the sign-up text appears in black and dark green font-color, the choices to "Keep me signed-in on this computer" and "Tailor Twitter based on my recent website visits," the Terms of Service, and the statement that, "… Others will be able to find you by name, username or email. Your email will not be shown publicly. You can change your privacy

settings at any time" appear in light grey. Further, the Terms of Service and note about public discoverability appear in a font-size smaller than the font-size used on the rest of the page. The note at the bottom is particularly difficult to read because it is light grey, in small text, and is set against a white background. While the locations where a user inputs text are clearly visible in larger font and using black text, in the locations where a user might learn about information flows, the text is sometimes made less visually prominent.

A second example of this pattern of font size and color can be seen within the new user tutorial. As can be seen in Figure 28, a description of how Twitter treats contact information within the "Find people you know" portion of the sign-up appears in smaller, light grey text.

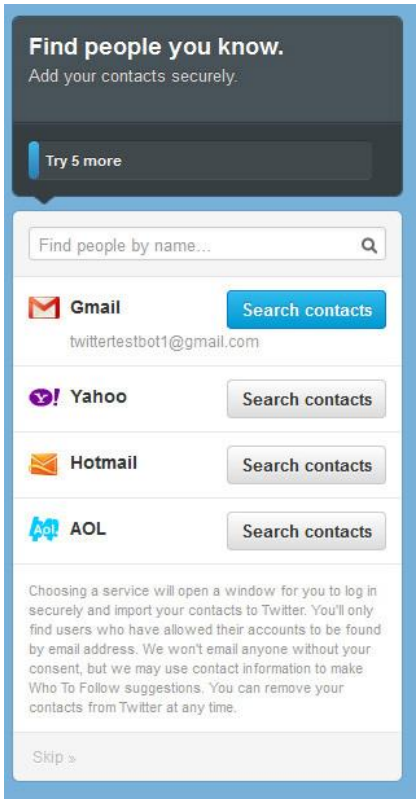


Figure 28. Screencapture of the "Find people you know" portion of the new-user orientation.

Additionally, in even lighter grey font in the bottom left of Figure 28 is the barely visible "skip" hyperlink. This is included because, according to the Privacy Policy, adding in contacts is actually an optional step, though the method through which one opts to not upload contacts is nearly hidden on the page. Here again, one can see that the locations where users are expected to upload information are clearly visible, yet the locations where a user might learn about how that information is treated, or learn that they can skip uploading this information are far less visible. As a result, this formatting functions as a potential hurdle in the development of users' *principles-knowledge* of information flows on the site.

Conclusion

In answer to the question, "how does the technological discourse surrounding the site created by Twitter's business purveyors describe information flows?" this analysis highlights a number different trends within the descriptions of information flows provided by Twitter, Inc. Many of these trends appear as though they may have significant consequences for users' abilities to develop *principles-knowledge* of different facets of the information flows on Twitter. As a result, these trends also have import to the broader question of the state of users' informational power in the user-Twitter relationship.

First, this chapter highlights a number of contextual factors that have shaped Twitter, Inc.'s choice of language within the corpus. These contextual factors include things such as the history of the documents themselves, the temporal proximity of Twitter, Inc.'s IPO, and the context of external governance and regional jurisdiction in which Twitter, Inc. operates. These factors are important to note because, despite the fact

that the documents in the corpus comprise the text that a user would encounter for the first-time in signing up for Twitter, each document has an independent production history and set of factors that has influenced the language that appears within. This is to say, the context in which Twitter, Inc. operates and in which the texts were produced has relevance for individual's ability to develop *principles-knowledge*, and hence the individual's informational power. As a result, the relationship between Twitter and user is also vicariously shaped by the relationship between Twitter and other actors, such as business partners, regional governments, the FTC, its investors, and even its own history.

Second, this chapter presents a number of trends that emerged from the analysis of the technological discourse surrounding the site created by Twitter's business purveyors. These trends were thematically organized around descriptions of who is involved with information flows, what information is flowing, when information flows, where information flows to and from, why information flows, and how information flows. Within each of these sections, the chapter elaborates on and provides examples of specific actors, types of information, places, justifications, and methods that involve different facets of data/metadata, algorithms, protocols, interfaces, defaults, informational content, users, governance, business practices, and platform ownership. While there are a number of locations in the corpus that provide specific and clear examples of how these facets of information flows are arranged on platform, they were greatly overshadowed by a large volume of vague, unspecific, implied, or hedged explanations.

Within the corpus, the kinds of data and metadata collected when a user interacts with Twitter are rarely explained, or explained in a way that leaves many of the details vague. Descriptions of protocols and algorithms tend to focus on explaining how

tweeting and following works, but not necessarily how other algorithms and protocols such as @replies or Trending work. As a result, users are oriented towards knowledge that facilitates network construction and information production, but not necessarily towards the details of how what they produce flows to others or becomes commodified by Twitter, Inc. or third-parties. Default settings are often described as something users can change at any time, but which cannot actually be set or viewed before or during registration. While a tutorial is given on the user-timeline interface as part of the orientation process, other aspects of the Twitter interfaces, such as the APIs, buttons, and widgets, go without definition or detailed explanation. As a result, users may be put at a disadvantage in their ability to understand information flows as they occur through non-timeline interfaces; for example, how visiting a site with a Twitter "Tweet This!" button relates to browsing behavior being reported back to Twitter, Inc. The corpus focuses heavily on the kinds (and social value) of informational content that Twitter gives users access to, and in particular, focuses on the "real-time" nature of both the information and the platform. However, the lifecycle of informational content is generally implied or described with caveats or ambiguity. Users on Twitter are repeatedly described as human, despite the fact that nearly 10% are not. Further, while there is some explanation of the different kinds of users on the platform such as registered users and unregistered users, there is little explanation about what kinds of differences there are between verified and unverified users. How information flows are tied to Twitter's business models and revenue generation methods also go generally undescribed or left implied. The fact that Twitter, Inc. sells access to the firehose of user-generated content goes entirely unmentioned. Twitter's business partners are hardly ever named explicitly (with the

notable exception of the mention of Google Analytics and the Library of Congress) and what these partners might do with user-generated data is generally not discussed. Ownership (and the fact that Twitter was to shortly become a publicly-traded company) is also not addressed anywhere in the corpus.

Third, there are an additional set of contextual factors that shape the readability and usability of the text. These additional presentation elements may also impact users' abilities to develop *principles-knowledge* about the information flows on Twitter. These included the use of hyperlinks in the corpus as a way of splitting up the text of the agreement between a user and Twitter, the fact that defaults and security/privacy settings cannot not be set during the sign-up process itself, and finally, the variations of text size and color within the corpus that made certain descriptions of information flows or the option to skip uploading certain information more difficult to perceive than other text.

With these observations about how the technological discourse surrounding the platform describes information flows now made, this chapter returns for one moment to a point about determinism. Despite the fact that Twitter, Inc. presents a particular picture of information flows on its platform, this does not mean that this picture unilaterally determines individuals' *principles-knowledge* of information flows on Twitter. Instead, this discourse must be thought of as one factor that can contribute to it. Readers of the corpus are free to ignore it, mistrust it, or interpret it in a number of ways. This is why the pairing of the discourse analysis with the user-knowledge study is so critical. While this discourse analysis has provided a number of observations about trends present in Twitter, Inc.'s descriptions of information flows on the Twitter platform, it is when these findings

are considered in tandem with users' knowledge that the findings truly have import towards addressing questions of informational power.

With these findings in mind, this dissertation now proceeds to its conclusion, which brings together the findings of this chapter and the previous chapter to address the primary research question: In the user-Twitter relationship, what is the state of Twitter users' informational power in regards to the informational flows of the platform?

Chapter 8: Conclusions about Informational Power in the User-Twitter Relationship

Introduction

To conclude this dissertation, this chapter addresses the third operationalized research question of the project (RQ3): In the user-Twitter relationship, what is the state of Twitter users' informational power in regards to the informational flows of the platform? To answer this question, this chapter draws together the findings from the fourth chapter's accounting of information flows on Twitter, the sixth chapter's survey and quantitative analysis of user knowledge of information flows, and the seventh chapter's critical discourse analysis of the language Twitter, Inc. presents during the new user registration process. After triangulating the findings of these chapters, this conclusion hypothesizes how the extant points of juncture and disjuncture can impact users' informational power and what the consequences of such impacts may be. Following this, the chapter suggests a number of remedies that could help address users' informational power moving forward. The chapter then offers a number of directions for future research before concluding.

Scope of the Findings and Generalizability

Before diving into the triangulation of how information flows on Twitter, user knowledge of information flows, and the results of the discourse analysis, it is first necessary to make a few remarks about the scope of the findings and their generalizability. As the quantitative analysis in this project uses a purposive and not random sample, the applicability of these findings to the larger population of concern—U.S. Twitter users over the age of 18—is unknown. Further study with true random

sampling is needed. However, the findings from Chapter 6 do offer an insight into where there may be points in users' *principles-knowledge* worth deeper investigation. Second, as discourse analysis relies on inductive reasoning and argument, the observations from Chapter 7 should not be read as making claims toward a single possible interpretation. Instead, the themes observed there should be understood as a particular, situated interpretation of the texts, where evidence from the application of a coding scheme is used to support the claims about prevalent themes and trends. Stemming from these limitations, this chapter's triangulation of the findings should also not be interpreted as making a claim towards a singular objective truth. Instead, this chapter should be read as an attempt at unpacking two potentially linked phenomena (knowledge and discourse) in an exploratory manner. It is in this spirit of exploration that this chapter will map the research findings: not from a position of absolute certainty about how well the findings generalize to all U.S. Twitter users and to all of the Twitter, Inc.'s discourse, but as a starting point in a conversation about the interrelations between information flow, user knowledge, technological discourse, and informational power.

Next, while this chapter is structuring the presentation of the results of the triangulation using a framework "juncture" and "disjuncture," this should not be interpreted as a strict binary relationship. Instead, these headers should be thought of as extremes of a continuum. Figure 29 provides a conceptual diagram of this continuum.

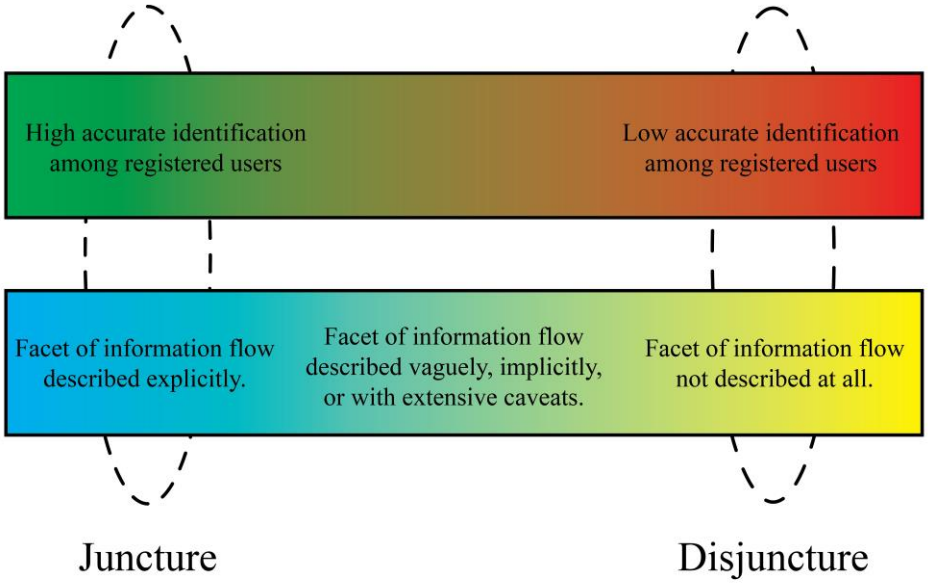


Figure 29. Conceptual mapping of juncture and disjuncture continuum.

Several of the information flow facets fall into a kind of middle ground between the two poles, where users do not demonstrate particularly high or particularly low rates of responding correctly to questions about a particular component of information flow, and/or where the corpus addresses a particular facet, but in a way that is vague, implicit, or heavily caveated.

As a result of a desire to focus on the clearest set of consequences, this chapter will not map the relationship between every single facet from Chapters 4, 6, and 7. There are simply too many facets of information flows, questions from the survey, and statements within Twitter's new-user orientation for that to be a practical exercise. Instead, in triangulating facets of information flows, users' principles-knowledge of the facets, and Twitter, Inc.'s discussion of the facets, this analysis inductively explores prevalent trends at the poles. It is in the mapping of the extremes that this chapter articulates the clearest set of consequences for users' informational power. The chapter does, however, give attention to divergent cases within the descriptions of juncture and

disjuncture, such as where user knowledge is particularly high despite a facet not being addressed within the corpus, and vice-versa.

Finally, while this chapter discusses a number of potential consequences that stem from particular states of users' informational power, these should be interpreted with a few caveats in mind. First, as *principles-knowledge* of the facets of information flow, and hence, informational power, varies by individual, so too does the potentiality of the consequences. Second, the potentiality of the consequences for users broadly is also dependent on whether or not the trends in user knowledge of information flows observed in the purposive sample did turn out to be representative of the wider population of U.S. Twitter users over the age of 18, *and* if this wider population interpreted the corpus exactly how it was interpreted in Chapter 7. Finally, as these findings are based on a snapshot of information flows, user-knowledge, and discourse taken at a particular moment in time, these consequences, of course, are respective to that temporal context. With those caveats in mind, this chapter now proceeds to describe the results of the triangulation of how information flows on Twitter, users' responses to survey questions about information flows on Twitter, and the discourse analysis of how Twitter, Inc. talks about information flows on Twitter as part of the new-user orientation process.

Juncture

There are several facets of information flows on Twitter where there appears to be alignment between users' understandings of the facet of information flow and how Twitter, Inc. talks about the facet as part of the new user orientation. A determination of alignment is dependent on two factors: 1) more than 50% of registered users having given an accurate response to the survey prompt about the facet, and 2) the interpretation that

the facet of information flow was described explicitly or in some detail in the corpus, based on the outcomes of the application of the coding scheme in Chapter 7. While the first component of this evaluation heuristic is easily quantified, the second is far more interpretive. As a result, the conclusions about juncture drawn here must be understood as an inductive interpretation.

It bears repeating this chapter is not arguing a kind of discursive determinism; i.e. that *because* Twitter, Inc. discusses these facets in some detail, users understand them. The language that Twitter, Inc. presents to users during the new-user registration process is merely one potential input from which a user can draw in building *principles-knowledge*. As discussed in Chapter 2, users can build their *principles-knowledge* via a number of different means and sources, such as by engaging other discourse (like newspaper articles about Twitter, talking with friends and family, consuming language that Twitter, Inc. has produced other than the new-user orientation), by directly interacting with and experiencing the technology, or watching others do the same. However, the discourse within the new-user registration process is unique among these ways of learning about Twitter. The new-user registration process is an important point where individuals are crossing a line between being a non-user or unregistered user, and becoming a registered user (assuming this is their first account). It is a formal moment of decision about technology adoption, or in the case of someone registering a second account, use re-evaluation. While the other ways that users may build up *principles-knowledge* are also important, what makes the new-user registration process unique is

that *all* registered users go through some form of it,[31] even if they choose not to read the language presented within it. In this way, it is perhaps one of the few experiences shared by nearly every registered Twitter user.

By placing the surveyed users' *principles-knowledge* of information flows next to this particular discourse, this section is not arguing direct causality, but is instead noting alignment. This alignment between users' demonstrated *principles-knowledge* and the discourse Twitter, Inc. presents is perhaps most apparent around facets of information flow on Twitter having to do with "first person informational production, consumption, and network building" and "advertisements as a business model."

First person informational production, consumption, and network building.

As the rank order chart of accurate responses to the survey prompts provided by registered users in Table 68 in Chapter 6 illustrates, more than 50% of registered users accurately responded to prompts having to do with the data, protocols, and algorithms associated with the production and consumption of content and network building on Twitter. This part of the "channel" of information flows appears particularly well understood by the sampled registered users. Specifically, prompts about how hashtags are used facilitate the organization of informational content on Twitter; whether "following" someone on Twitter means that you are subscribing to their Tweets as a follower and their updates will appear in your Home tab; what makes a Tweet an @reply; that having a protected account in part means that others wishing to follow you must be approved by you; how many characters Tweets are limited to; the fact that the trending topics algorithm displays hashtags that are popularly in use in the short-term; the fact that,

[31] Although once Twitter began to allow registration through its official mobile app, the experience was no longer consistent across all users, although it is very similar on the mobile app. The language on the mobile app is an area deserving of exploration in future research.

unless a user makes changes to the default settings, Tweets are automatically set to public; and the fact that users can share geo-location information in Tweets, all had accurate identification rate in excess of 50%.

Many—but not all—of these facets were recurrently discussed within the corpus. Further, the corpus explains these facets in fairly plain and clear language. For example, as part of the new-user tutorial, the "Twitter Teacher" explicitly explains to users that Tweets are limited to 140-characters. The "@Home" interface also provides a running count of characters used when composing a new Tweet. The new-user tutorial shows registering users how to follow other users and asks them to follow five other users to get started, and the Twitter Teacher demonstrates how the Tweets from the newly followed users populate into the timeline within the user interface. The corpus describes the public status of Tweets in the first line of the Privacy Policy, when it states "Any registered user can send a Tweet, which is a message of 140 characters or less that is *public by default* and can include other content like photos, videos, and links to other websites [emphasis added]." The corpus addresses the fact that users can choose to upload geo-location information along with Tweets when the Privacy Policy states:

> You may choose to publish your location in your Tweets and in your Twitter profile… You may also tell us your location when you set your trend location on Twitter.com or enable your computer or mobile device to send us location information.

The corpus does so again on the "Security and Privacy" settings page, where there is a clearly visible check-box with the statement, "Add a location to my Tweets." The corpus also explicitly addresses protected accounts on Twitter in the "Security and Privacy" settings page, which explains under the option to "Protect my Tweets," "If selected, only those you approve will receive your Tweets." Thus, for these facets of information flows

on the Twitter platform, there appears to be juncture between users' *principles-knowledge* and the discourse provided by Twitter, Inc. provides.

There were, however, a number of components of information production where a majority of registered users gave accurate responses, but where the corpus either did not directly or only implicitly discussed the facet. For example, 95% of registered users were able to identify the statement "Including a hashtag (the '#' symbol) in front of a word is a way of marking keywords or topics in a Tweet and is sometimes used by users as a way to categorize messages" as correct; however, the corpus only vaguely and implicitly addresses hashtags. Within the corpus, hashtags only appear in the "#Discover" tab and within the "Trending Topics" area of the user interface. These areas do not include explanations of what hashtags are or how they work. However, there are many contextual clues which suggest users insert hashtags into Tweets as a way of marking topics within Tweets, which thereby links the messages. Hashtags have also seemed to enter the popular vernacular and are not exclusive to the Twitter platform, which may in part help explain the high percentage of users with accurate responses.

Similarly, 51.9% of the sampled registered users gave accurate responses to the correct statement: "Twitter's trending topic algorithm identifies topics that are immediately popular, rather than topics that have been popular for a while or on a daily basis." The phrase "trending topics" appears in multiple locations within the corpus; however, the corpus does not explain in detail how trending topics are determined or discuss the fact that the algorithm excludes topics that have been popular for some time. There are, however, context clues present that might allow a user to piece together a rough idea of how the algorithm functions. The phrase "trending topics," may itself be

rather self-explanatory and suggest the contemporary focus of the algorithm. It is also possible that Twitter, Inc.'s recurrent descriptions of Twitter as a real-time platform contributes to the high volume of users who gave accurate responses. Interestingly, when the survey presented respondents with the incorrect statement, "All users see the same trending topics," only 45.3% of registered users were able to accurately identify this as an incorrect statement, 39.7% of registered users indicated that they were not sure, and 14.5% inaccurately identified the statement as correct. This suggests that more registered users may be uncertain or outright incorrect in their *principles-knowledge* regarding what other users see from the trending topics algorithm than how the algorithm determines which topics appear. Nowhere does the corpus explicitly address the fact that different users see different trends, and hence, no opportunity within the new user registration process for this (lack of/mis) understanding to be countered.

There was also a high rate of accurate identification of the statement, "Including the "@" symbol and another user's Twitter username (such as "@PBS") at the beginning of a Tweet means that Twitter will treat that as a "reply" to that user" (72.3% accurate identification). This is interesting because the corpus does not directly describe what makes a Tweet an @reply. However, the @reply button is built into the user interface itself, is clearly visible when Tweets appear in the user's timeline, and pressing the button automates the insertion of the @username text at the beginning of the Tweet. This suggests that users may develop this part of their *principles-knowledge* more through use or means other than the new user-orientation. This finding can also be juxtaposed with the finding that only 25% of registered users gave an accurate response to the correct statement (Q18): "A Twitter user will only see another user's @replies in their home

timeline if they are following both the sender and recipient of the @reply." Fifty percent of users indicated that they didn't know whether the statement was correct or incorrect and 25% gave an inaccurate response. This suggests that while registered users may have strong *principles-knowledge* about what makes a Tweet an @reply, they may not be fully aware of how @replies do or do not flow the timelines of their followers. There is also nothing in the corpus that addresses how @replies do or do not flow to the timelines of others.

Looking across these findings, there appear to be several facets of information production, consumption, and network building where a majority of registered users demonstrated accurate *principles-knowledge* and where the corpus seems to support the development of this *principles-knowledge*. While this section presents such findings as "junctures," these overlaps do each have their own individual consistency. For example, while almost 95% of registered users gave accurate responses to the statement about what hashtags do, just over 50% accurately responded that users can share geo-location information within Tweets. The corpus describes some facets (such as the 140-charachter length) recurrently, whereas others are mentioned only once.

As discussed further in the "Disjuncture" section, this group of findings is also interesting because it predominantly relates to the first-person experience of the information production, consumption, and networking facets of the channel of information flows. This stands in contrast to the disjunctures present in users' understandings of how the content users create (such as @replies) flow to others, or what others see from the trending topics algorithm.

Advertisements as a business models.

Registered users had relatively high rates of accurately identifying promoted Tweets, promoted trends, and promoted accounts as ways that Twitter generates revenue (at rates of 73.2%, 56.2%, and 47.2%, respectively). Although Twitter, Inc. never explicitly addresses the commercial nature of the advertisements within the corpus (although this may somewhat self-evident), the corpus does mention multiple times that there is advertising on the site and that Twitter works with a variety of third-party advertisers. Conversely, only 21.7% of registered users correctly identified the fact that Twitter, Inc. generates revenue through the sale of access to user-generated content through via the firehose API. In fact, there was no statistically significant relationship between user-status and response to this prompt. This stands in contrast to the other three revenue generation models, which did have a statistically significant relationship between user-status and response. The sale of the real-time firehose as a business model is never addressed clearly and explicitly within the corpus, and as a result, there is little in the corpus that would seem to support users building that aspect of their *principles-knowledge* of information flows on Twitter.

What this suggests is that many users conceptualize Twitter's revenue generation as advertising based, but do not necessarily see their own information production as something specifically commodified by Twitter. Nor does Twitter, Inc. present its business models as part of the new-user orientation process in a way that would contradict that conceptualization. Twitter, Inc.'s Certified Products Partners, many of who resell Tweets, similarly go unnamed and their commercialization of user-generated content goes unmentioned in the corpus. Given this, perhaps it should not come as a surprise that when registered users were presented with the incorrect statement:

"Twitter's 'Certified Products' partners are all prohibited from reselling historical Twitter data, such as old Tweets, to third-parties," over 80% of registered users indicated that they did not know whether the statement was correct or incorrect. While there is juncture regarding advertisements as business models, as discussed further in the next section, there is disjuncture around the commodification of the user-generated content.

Disjuncture

Registered users demonstrated high rates of either uncertainty or outright misunderstanding about several facets of information flows on Twitter. In many of these cases, the corpus provides either unclear descriptions of these facets or does not address them entirely. However, this is not to say that *because* Twitter, Inc. does not address these facets or does not do so in detail that users do not understand them. Instead, this is to say that, in many cases, there are not clear statements in Twitter, Inc.'s new-user registration process that could help dispel these misunderstandings or uncertainties.

The determination of "disjuncture" around a particular facet is based two factors: 1) less than 35% of registered users having accurately responded to the statement about the information flow facet, and 2) the interpretation that the facet was either not described in the corpus, or was described an implicit or unclear manner, based on outcomes of the coding process. The disjuncture among registered users' surveyed *principles-knowledge* and the discourse presented by Twitter, Inc. in the new user registration process is perhaps most visible around facets of information flow on Twitter having to do with "what other users see or send," "what data Twitter collects," "how Twitter makes information accessible," and "what happens to user-generated information."

What other users can see or send.

As briefly introduced in the "First person informational production, consumption, and network building" section, registered users who participated in the survey frequently had high rates of giving inaccurate and/or "don't know" responses to prompts having to do with what information other Twitter users can see or send. These facets have to do with the "channel" part of information flows, but not the part that a user directly experiences (it is instead, the "receiving" end of the information flow). Registered users performed poorly on questions about whether Twitter ever withholds Tweets in certain countries; the visibility of @Replies on followers' timelines; whether or not @replies from people you do not follow are shown on your timeline; whether protected Tweets can be seen by non-followers if they are @mentioned within a Tweet; if you have a "protected" account on Twitter, whether your Tweets are visible to the followers of your followers; whether unregistered users can access search; whether information about the number of Tweets, number of photos, number of followers and followees and number of favorites of users from protected accounts is publicly accessible; and whether any user can direct message any other user by default each had accurate response rates south of 35%.

In many of these cases, there are not statements within the corpus that could help dispel users of their misunderstanding or lack of understanding. The corpus never explicitly addresses the fact that Twitter withholds Tweets in certain countries. Instead, a reader must recognize this as an implication of Twitter's statement that it complies with the local laws of the countries it operates in. The corpus does not explain how @replies do and do not propagate to followers' timelines, if @mentions that come from protected accounts can be seen by non-followers, or if @replies from users who do not follow you

can be seen on your timeline. It also does not explain how the direct message defaults are arranged on the platform and whether anyone on the site can direct message anyone else.

Similarly, a description of whether information about the number of Tweets, number of photos, number of followers and followees, and number of favorites of users with protected accounts remains publicly accessible is present in the corpus, although this aspect is described in an implicit manner. The Privacy Policy states:

> Our Services are primarily designed to help you share information with the world. Most of the information you provide us is information you are asking us to make public. This includes not only the messages you Tweet and the metadata provided with Tweets, such as when you Tweeted, but also the lists you create, the people you follow, the Tweets you mark as favorites or Retweet, and *many other bits of information that result from your use of the Services*. Our default is almost always to make the information you provide public for as long as you do not delete it from Twitter, but we generally give you settings to make the information more private if you want. Your public information is broadly and instantly disseminated. [emphasis added]

A reader must recognize that "many other bits of information" includes counts of numbers of Tweets, numbers of photos, numbers of followers and followees, and numbers of favorites. However, this statement does not clearly and explicitly indicate that even if you choose to protect your account, some metadata about your account remains publicly accessible, because this metadata is not clearly defined as "public information" (though it is also not defined inversely as private information). In the only other location in the corpus where account protections are discussed, the "Security and Privacy" settings page of the interface, next to the subheader "Tweet privacy," there is checkbox option to "Protect my Tweets." Under this appears a statement, "If selected, only those you approve will receive your Tweets. Your future Tweets will not be available publicly. Tweets posted previously may still be publicly visible in some places." The statement

remains entirely silent on whether the number of photos, number of followers and followees, and number of favorites are protected if a user chooses to protect their Tweets.

While the fact that the corpus does not address or ambiguously addresses many information flow facets having to do with what other users can send and see might seem a likely justification for users' lack of *principles-knowledge*, an entirely a causal relationship seems unlikely. This is because, in a handful of locations, users demonstrated poor *principles-knowledge* about some facets of information flow that the corpus discusses explicitly. For example, the statement on the "Security and Privacy" settings page for "Protect my Tweets" states: "If selected, only those you approve will receive your Tweets." This should suggest to a reader that only those you approve will receive your Tweets, and hence, followers of followers would not be eligible to view one's protected Tweets. Similarly, 33.3% of registered users gave an inaccurate response and 46.2% of registered users gave a "don't know" response when presented with the incorrect statement: "Unregistered visitors to Twitter.com can still view publicly created Tweets but cannot use the "search" feature of the website." Yet, the Privacy Policy addresses this facet fairly explicitly when it states, "Some Services, such as search, public user profiles and viewing lists, do not require registration." While this statement does not use the phrase "unregistered users," it explicitly names search as an example of a service that can be accessed without needing registration. It should be said, however, that this is the only location in the corpus that explicitly addresses unregistered users' access.

What data Twitter receives / collects.

A large percentage of registered users gave inaccurate or "don't know" responses to prompts about what data Twitter collects from users during their use of the platform.

This involves both knowledge of the "channel" of information flow and how Twitter, Inc. functions as a user and receiver of information. For example, questions about whether Twitter automatically collects GPS information with Tweets by default and whether Twitter receives data about visits to sites that have embedded widgets and buttons both had accurate responses rates less than 35%.

Despite users having relatively high rates of ambiguity in their understandings of these facets, they are both addressed in corpus, although with questionable clarity. The corpus addresses how and when GPS information is collected by Twitter in the Privacy Policy when it states:

> You may choose to publish your location in your Tweets and in your Twitter profile. You may also tell us your location when you set your trend location on Twitter.com or enable your computer or mobile device to send us location information. You can set your Tweet location preferences in your account settings ["account settings" is hyperlinked to account settings page which requires a user to be logged in and for registration to be complete] and learn more about this feature here ["here" is hyperlinked to FAQ on adding locations to Tweets].

While this language may not be the clearest, the use of the term "may choose" in the passage indicates that a user optionally provides this information. Further, the statement suggests that a user must enable their computer or mobile device to send location information, making it appear as though enabling is an active decision on the part of the user.

The fact that Twitter, Inc. receives the URLs of user site visits to third-party websites with buttons and widgets is addressed in the corpus, though the explanation falls over multiple paragraphs and sections. In its section on "Log Data," The Privacy Policy states:

> Our servers automatically record information ("Log Data") created by your use of the Services. Log Data may include information such as your IP address, browser

> type, operating system, *the referring web page*, *pages visited*, location, your mobile carrier, device and application IDs, search terms, and cookie information. We receive Log Data when you interact with our Services, for example, when you visit our websites, sign into our Services, interact with our email notifications, use your Twitter account to authenticate to a third-party website or application, *or visit a third-party website that includes a Twitter button or widget* [emphasis added].

In its section on "Widget Data," the Privacy Policy expands, stating:

> We may tailor content for you based on your visits to third-party websites that integrate Twitter buttons or widgets. When these websites first load our buttons or widgets for display, *we receive Log Data, including the web page you visited and a cookie that identifies your browser* ("Widget Data") [emphasis added].

However, as noted in Chapter 7, nowhere in the corpus are readers given an example or definition of what a widget or button is or looks like. While individuals with a higher level of digital literacy might be able to recognize what a widget or button is, those without this prior knowledge may be lost. So while both facets are described in the corpus, there are barriers in this description of information flow that may inhibit the easy development of *principles-knowledge* around these facets.

Twitter's APIs.

A high percentage of registered users gave "don't know" responses to prompts having to do with Twitter's APIs. This involves the structure of the "channel," (as the APIs function as an interface), who the users (receivers) of this channel are, the business models involved in the arrangements of information flows through this interface, and how the use of information flows from this part of the channel are governed. For example, questions about whether Twitter offers interfaces for programmers and application developers, which can be used to make applications, websites, widgets, and other projects that interact with Twitter; whether Twitter offers a search interface to programmers that allows them to search for Tweets by latitude, longitude and radius area,

and will attempt to find Tweets created by users whose profile location matches the search parameters; and whether users of Twitter's APIs are forbidden from selling, renting, leasing, or redistributing access to the Twitter API or Twitter Content to any third party without prior written approval from Twitter all had accurate response rates less than 35% (and additionally "don't know" response rates in excess of 65%).

As observed in Chapter 7, while the corpus addresses the fact that Twitter offers APIs, they are discussed without an initial explanation of what they are. For example, the first mention of the API in the Terms of Service states, "These Terms of Service ("Terms") govern your access to and use of the services, including our various websites, SMS, APIs, email notifications, applications, buttons, and widgets…" A reader who does not already know what an API is must guess that an API is some way of interacting with the Twitter platform, based on its inclusion in a list of other ways of interacting with the Twitter platform. Later, the ToS provides slightly more context when it states:

> Except as permitted through the Services, these Terms, or the terms provided on dev.twitter.com, you have to use the Twitter API if you want to reproduce, modify, create derivative works, distribute, sell, transfer, publicly display, publicly perform, transmit, or otherwise use the Content or Services.

Underneath this statement, the document offers a tip: "We encourage and permit broad re-use of Content. The Twitter API exists to enable this." There is no discussion of *who* might use the APIs or might be interested in the re-use of Content, how the APIs exactly enable re-use, or what the APIs do and do not allow for in terms of gathering and using information from Twitter.

The fact that the APIs can be used to search for Tweets by geo-location included in Tweets or the geo-location of users' profiles is not explained in the corpus. The only

place that the corpus comes close to detailing this facet is in the "Location Information" section of the Privacy Policy, in which the document states:

> We may use and store information about your location to provide features of our Services, such as Tweeting with your location, and to improve and customize the Services, for example, with more relevant content like local trends, stories, ads, and suggestions for people to follow.

While the search functions within the APIs are included as part of the "Services," this is left implicit.

Similarly, there is not a clear discussion in the corpus of whether users of Twitter's APIs must seek permission from Twitter to sell, rent, lease, or redistribute access to the Twitter API or Twitter Content. The ToS does state that, "…you have to use the Twitter API if you want to reproduce, modify, create derivative works, *distribute*, *sell*, transfer, publicly display, publicly perform, transmit, or otherwise use the Content or Services [emphasis added]." However, the fact that this does not mean that API users have carte blanche to do all of those things without first getting prior approval is not actually spelled out by this statement. The rules that govern acceptable and unacceptable uses of the Twitter APIs are detailed elsewhere on Twitter. As a result, users are at a disadvantage in their ability to develop *principles-knowledge* of what API users can and cannot do with the content users generate, and how exactly API users are governed, if user knowledge was based solely on the new-user registration process.

What happens to Tweets.

A large percentage of registered users gave inaccurate or "don't know" responses to statements having to do with what happens to Tweets in the long-term. This relates to who acts as a receiver of information flows and the governance of information flow. For example, questions about whether old Tweets are automatically deleted from Twitter's

servers after 2 years; whether Twitter's ‘Certified Products’ partners are all prohibited from reselling historical Twitter data, such as old Tweets, to third-parties; and whether Tweets created by users with public accounts are given to the Library of Congress for archival all garnered “don’t know” and inaccurate responses in excess of 65%.

As described in Chapter 7, the corpus does not focus much attention on the lifecycle of Tweets, only discussing it in a few locations. Instead, the corpus focuses much more heavily on positioning Tweets and the Twitter platform as real-time and immediate. When the corpus does discuss the lifecycle of Tweets, the lifecycle laid out is hedged. The Privacy Policy suggests that Twitter, Inc. stores Tweets indefinitely when it states: “Our default is almost always to make the information you provide public *for as long as you do not delete it* from Twitter.” However, the ToS throws this into question when it states, “We … retain the right to create limits on use and storage at our sole discretion at any time without prior notice to you.” Together, these two statements suggest that Tweets are stored until a user asks to delete them, but that Twitter, Inc. may also independently make the decision to limit the time of storage for Tweets and may not tell users if and when it does so. As a result, the corpus describes a lifecycle for Tweets, but one marked by some level of ambiguity.

While the corpus does state that Twitter, Inc. works with a number of business partners, it never spells out who these business partners are, what these business partners do, and how these business partners use user-generated content. More narrowly, the corpus never once mentions Twitter’s “Certified Products” program, nor the fact that some of Twitter’s Certified Products resell user-generated content. Instead, this is left

implied by the ToS as part of the explanation of the rights that users grant Twitter, Inc. regarding the informational content they produce:

> You agree that this license includes the right for Twitter to provide, promote, and improve the Services and to make Content submitted to or through the Services available to other companies, organizations or individuals who partner with Twitter for the syndication, broadcast, distribution or publication of such Content on other media and services, subject to our terms and conditions for such Content use.

While this passage suggests that user-generated content may be syndicated, broadcast, distributed, or published by companies that partner with Twitter, Inc., the statement does not spell out the commercial nature of some of this syndication, broadcast, distribution and publishing.

While the lifecycle of Tweets and the fact that older data is resold by Twitter's business partners are not described with exceptional clarity by the corpus, the Privacy Policy does explicitly name the Library of Congress as a recipient of Tweets and indicates that they archive Tweets for historical purposes. The Privacy Policy states:

> For instance, your public user profile information and public Tweets may be searchable by search engines and are immediately delivered via SMS and our APIs to a wide range of users and services, with one example being the United States Library of Congress, which archives Tweets for historical purposes.

However, this is the only location where the corpus ever mentions the Library's archival. Further, the Privacy Policy provides no information about how the Library of Congress does or does not make the archive available to others, whether or not Tweets that are deleted from Twitter are also deleted from the LoC archive, or if an individual chooses to make their account protected exactly what information the LoC still receives (such as whether or not profile information, meta account information such as number of followers, number of Tweets, number of favorites and so on are archived). Thus, the

corpus facilitates the development of *principles-knowledge* regarding the fact the archiving takes place, but does not detail how an individual can or cannot control how their information is stored by deleting content, the full picture of what user-information is stored, or whether individuals do or do not they have the ability to "opt-out" of such archiving.

Twitter as a business.

A high percentage of registered users gave inaccurate or "don't know" answers in response to prompts having to do with facts about Twitter as a business, including its ownership and some of its business models. For example, questions about whether Twitter generates revenue by selling access to the full stream of real-time Tweets created by users to third-parties and whether Tim Cook is the current CEO of Twitter both garnered accurate response rates lower than 35%.

The corpus addresses neither of these facets. The fact that Twitter sells access to the firehose API must be inferred from the Terms of Service, which states that when a user signs up to Twitter:

> By submitting, posting or displaying Content on or through the Services, you grant us a worldwide, non-exclusive, royalty-free license (with the right to sublicense) to use, copy, reproduce, process, adapt, modify, publish, transmit, display and distribute such Content in any and all media or distribution methods (now known or later developed)…You agree that this license includes the right for Twitter to provide, promote, and improve the Services and to make Content submitted to or through the Services available to other companies, organizations or individuals who partner with Twitter for the syndication, broadcast, distribution or publication of such Content on other media and services, subject to our terms and conditions for such Content use.

The Terms of Service provides a shorthand "tip" of this passage, that, "This license is you authorizing us to make your Tweets available to the rest of the world and to let others do the same." The fact that Twitter sells access to a real-time stream of all Tweets is not

spelled out within this language. Similarly, the texts never address ownership or management of Twitter, Inc. nor the fact that Twitter was to shortly become a publicly-traded company.

Summarizing the State of Twitter Users' Informational Power

At the heart of this dissertation lies the premise that there is a connection between an individual's *principles-knowledge* of information flows on the Twitter platform and their relative informational power, and that this informational power can be shaped in part (although not deterministically nor entirely) by the individual's internalization of the discourse that addresses the information flows of the platform in the new user registration process. Now that this chapter has mapped some of the prominent junctures and disjunctures between the way information flows on Twitter, the way Twitter, Inc. talks about information flows on Twitter in the new user registration process, and the ways that registered users from the purposefully sampled population maintain *principles-knowledge* about information flows on Twitter, this chapter now turns to reconnect to the conversation about power. In doing so, this section will address some of the potential consequences of these junctures and disjunctures for individual users.

Rather than presenting a review of every single facet of information flows on Twitter, users' understandings of that flow and the way that Twitter, Inc. describes that flow, along with the potential hypothetical consequences for individual power that stem from the juncture/disjuncture among the three, this section will instead present a selection of thematic trends from the results of the triangulation, exploring some of the potential consequences of the points of overlap and disjoint. This review will also situate the consequences alongside some of the extant body of research introduced in Chapter 3. The

consequences for user informational power are described here under the headings of "limits in relation to commodification," "limits in relation to long-term information flows," and "information flow solipsism."

Limits in relation to commodification.

As noted at the beginning of this chapter, a higher percentage of registered users accurately identified promoted Tweets, promoted trends, and promoted accounts as ways that Twitter generates revenue than in comparison to the percentage that accurately identified the sale of access to user-generated content via the firehose API. Further, fewer than 10% of registered users gave an accurate response to a prompt about the fact that Twitter's Certified Products resell user-generated content. Respondents demonstrated weak *principles-knowledge* about how the socio-economic dimension of business models are part of information flow, how Twitter's business partners act as a receiver of informational content, and how those business partners subsequently construct their own for profit information flows using user-generated informational content. What this suggests is that many users conceptualize Twitter's business models as advertising based, but fewer recognize how their own information production is commodified by Twitter or Twitter's partners. The corpus does little to dispel this state of uncertainty or outright misunderstanding among users, as it makes regular mention of the fact that there is advertising on the site, but never explicitly describes the sale of access to informational content via the firehose API interface or that Twitter's "Certified Products" may resell user-generated content.

From this, it appears that users do not have much in the way of *principles-knowledge*, and hence, informational power, in regards to the fact that Twitter, Inc. and

Twitter's business partners commodify user-generated content. Twitter, Inc. does not say much in its new user registration process that would help a user develop this *principles-knowledge*. While it is difficult to argue that Twitter, Inc. *intentionally* limits the development of users' *principles-knowledge*, and hence, informational power about this particular facet, they offer little in the new user orientation that would help registering users build this informational power. Thus, this dissertation argues that, in the user-Twitter relationship, many users have a weak base of informational power in relation to the commodification of the information they generate.

This finding appears to run parallel to Langlois, McKelvey, Elmer and Werbin's (2009) argument that a common characteristic of commercial Web 2.0 platforms is that user-generated content commodification is often rendered invisible to users through the technical structuring of these spaces. The findings of this dissertation build out the work of Langlois et al. by suggesting that, in addition to the technical structuring (code politics), there is a discursive component to this process; the commodification of user-generated content is not explicitly and clearly described as part of Twitter, Inc.'s projection of information flows of the platform. By mapping the connections (and more frequently, the disconnections) between user knowledge of information flow commodification, the way that information flow commodification is described within the platform, and the actual information flow commodification, this work provides a basis of empirical evidence to support Langlois et al.'s claim that, "Web 2.0 spaces serve to establish the conditions within which content can be produced and shared and where the sphere of agency of users can be defined" (para. 15). Langlois et al., suggest that Web 2.0 organizations obfuscate the commodification through code politics in order to reduce

"resistance" among users and, thus, be better able to extract value from users' labor. In doing so, the organizations shape the horizons of possibility for the individual. While this dissertation cannot speak to the intentions of Twitter, Inc., it can speak to the potential outcomes for individual's informational power based on the evidence observed in the mapping of juncture and disjuncture.

There are a number of fields of action that a user might choose to pursue with a more robustly developed base of *principles-knowledge* about the commodification practices around user-generated content. On the basis of this informational power, a newly registering user might choose to abandon the account sign-up process, or might choose to continue with registration but be careful about the kinds of information they provide to Twitter. For example, they might choose a false name, use temporary e-mail address, or may otherwise choose to obscure their identity to avoid the linking of their identifiers to content they produce. With knowledge of these commodification practices, a user might petition Twitter in order to put pressure on them to change the practices, similar to how Facebook users petitioned Facebook to end the Beacon program after the details of how it worked became public. They might change their use habits by doing things such as not mentioning brands or not using specific words in order to make the informational content they produce less valuable to those who would buy access to the real-time stream of Twitter data. Or, somewhat conversely, they might make the informed decision that they are entirely comfortable with the commodification practices on Twitter and continue using the service with no changes to their behavior. They might decide that they want to take advantage of the firehose API and purchase access to it. They may decide that they are interested in purchasing data from some of Twitter's

"Certified Products" in order to do historical research. These possibilities are however, closed off, when users do not have informational-power in regards to commercialization of information flows necessary to enter into these fields of action.

Limits in relation to long-term lifecycle of information.

Similar to facets of information flow having to do with commodification, users also demonstrated relatively weaker *principles-knowledge* regarding the storage and use of older data. As noted in the beginning of this chapter, registered users had high rates of giving "don't know" responses to questions about whether Twitter's Certified Products resell old Tweets, whether Tweets are archived by the Library of Congress, and whether Twitter deletes old Tweets after two years. This suggests that while the immediate arrangement of facets of information flow (particularly around the information channel) might be well understood by users, how information flows are arranged in the long-term is less well understood. However, the way the corpus describes these facets varies more than in its discussion of the commodification of user-generated content. The corpus does not explicitly address the fact Twitter's Certified Products resell old Tweets. The corpus indicates that Twitter generally makes Tweets available unless they are deleted, but does create some ambiguity when the ToS notes that Twitter reserves the right to limit its storage without prior notice. The Library of Congress's archival of Tweets is explicitly mentioned in the Privacy Policy (albeit only once). However, overall, the corpus focuses more on positioning Twitter and the informational content on Twitter as being "real-time" than discussing the long-term lifecycle, transmission, commodification, and storage of user-generated content. These findings suggest a complex picture for users'

informational power in relation to the long-term or historical nature of Tweets in the user-Twitter relationship.

Based on the observed trends, one could argue that users maintain a relatively weak base of informational power in relation to the long-term of Tweets. However, the corpus may not be entirely culpable in this state of user power. While the corpus does not lavish attention on the long-term of information uploaded to Twitter in the way it provides descriptions of content on Twitter as about the real-time, it does at least touch on some aspects of the data lifecycle, albeit sometimes quite briefly. While users may be able to build some principles-knowledge of the long-term of Tweets from the corpus, this *principles-knowledge* seems unlikely to be as robust as the *principles-knowledge* about the real-time nature of the platform. Hence, it appears that users' informational power is more oriented towards the real-time production and consumption than in relation to applications having to do with the long-term of the content they generate.

One might ask: what would an individual with a more robustly developed base of informational power in regards to these long-term facets of information flow do differently? There are a number of fields of action that such *principles-knowledge* might open up. For instance, an individual who has the *principles-knowledge* that the Library of Congress archives Twitter may decide that they want to investigate how this arrangement came to be. In light of such *principles-knowledge*, that individual might consider filing a Freedom of Information Act request to the Library of Congress asking them for copies of their business agreements with Twitter. Individuals with such knowledge may want to protest such an information flow by writing to their Congressperson, asking them to intervene and stop the archiving. They may be individuals in positions of structural power

that could take advantage of the archive. An individual may want to contact the Library of Congress in order to see if they can use the archive for historical research, to gain a better understanding of how Twitter users responded to various world events. They may want to stop using Twitter altogether to protest such archiving. This informational power, in effect, opens up a number of potential fields of actions from which the individual can then choose to or not to pursue particular paths. These paths importantly include the non-use or non-adoption of the technology.

At a conference where I was recently presenting on some of the initial analysis of Chapter 6, an individual told me that, in light of knowing about the archiving by the Library of Congress, she began keeping two accounts; one, a protected account, and the other, a publicly accessible account that she purposively uses to tweet things that she wants to be included in the Library of Congress's archive. Her *principles-knowledge* of this flow serves as a base of informational power from which she makes decisions about which account to use, and whether or not she wants to contribute to the "historical record."

Alternatively, with detailed knowledge about how historical data is commodified and stored indefinitely, users might seek out a service like TweetDeleter[32] or TweetEraser,[33] which allow a user to delete Tweets based on the year the Tweets were written, based on specific content the Tweet contains, or based on the age of the Tweet. In a 2015 article on the tech news site *Fusion*, author Kevin Roose interviewed a number of individuals (including a former Twitter employee) who have chosen to use scripts or

[32] http://www.tweetdeleter.com/en
[33] http://www.tweeteraser.com/

other programs such as those mentioned above that allow them to control the lifecycle of Tweets through timed-deletion. Roose writes of one such interview:

> Josh Miller, a product manager at Facebook, wrote a piece of code that deleted his tweets after seven days. He frames his tweet-deleting as a decision to make Twitter more like other forms of conversation. "My opinions aren't permanent in my head (I often change my mind over time), and they're not permanent when shared around the dinner table (nobody is recording our conversations)," Miller wrote in an e-mail. "So it just doesn't make sense to me that they would be permanent online." (Roose, 2015, para. 9)

Interestingly, Roose also connects the issue of lifecycle to commodification in a discussion with a former Twitter employee:

> Social media companies, predictably, aren't thrilled with the idea of users mass-deleting their posts. With the exception of Snapchat and a handful of other apps, these sites are built on the idea of lasting data. Our posts are meant to stay up indefinitely, each one a piece of an ever-expanding mosaic of our desires, tastes, and preferences. If all of their users auto-deleted their old posts, Facebook, Twitter, and other social networks would have a hard time constructing the user profiles that are used as bait for advertisers. The entire business model might collapse. "If anyone ever seriously proposed [a tweet auto-deleter], they were quickly shot down," Sloan recalls of his days working at Twitter. "When you have a huge, deep corpus like this, you can do interesting stuff with it." (A Twitter spokeswoman declined to comment.) (Roose, 2015, para. 16)

Perhaps the economic justification given also explains why such little attention is given to the long-term of user-generated content within the corpus.

With greater principles-knowledge of the long-term of information flows on the platform, users might choose what they tweet more carefully, might stop using the service altogether, might try to exercise greater control over the lifecycle through the use of third-party services or scripts that automate deletion, or might consciously decide to that they are entirely comfortable with the current arrangement and continue using the service just as they had been before. Again though, these possibilities are closed off when users do not have informational power to enter these fields of action.

Information flow solipsism.

To summarize the major observations about users' *principles-knowledge* and informational power that have been made in this chapter so far, the registered users sampled have strong *principles-knowledge* of data, protocols, and algorithms (pieces of the information channel) associated with the first person-experience of production and consumption of information on Twitter, following protocols, and the fact that Twitter, Inc.'s business practices include making money through advertising. The corpus generally supports the construction of knowledge around these facets. The registered users sampled demonstrated relatively weaker *principles-knowledge* about what information other users can see from the receiving end of the informational channel (such as whether everyone sees the same trends; if followers see @replies; if Tweets are ever restricted based on location); about the kinds of data and metadata that Twitter receives and collects about users; about the API interfaces, who receives data from them, and the governance of the use of those interfaces; about Twitter, Inc.'s business practices of commodifying user-generated content through the sale of access to the firehose API and Twitter, Inc.'s business partners' resale of user-generated informational content; and about the arrangements of information flows beyond the real-time, such as the archiving of Tweets by the Library of Congress and Twitter's own Tweet storage practices. In many (but not all) of these cases, the corpus stays relatively silent or ambiguous about the specifics of these facets of information flow, at least in comparison to the attention given to facets having to do with users' information content production, consumption, networking, and the real-time nature of the platform. Taken together, these findings suggest that users' *principles-knowledge*, and hence, informational power, is highly

geared towards fields of action that include producing and consuming real-time information, networking, and advertising, and not as geared towards fields of action having to do with considering information flows from other users' perspectives, having to do with how information flows beyond the immediate or real-time configuration, with how parties beyond other registered users function as receivers, gaining access to and using the information users produce, with Twitter's business practices beyond advertising, or with how the information users create flows within an "ecosystem" larger their than own network. I suggest this state of individuals' *principles-knowledge*, and hence, informational power, can be described best as a state of "information flow solipsism." I further posit that the corpus, overall, does little to dispel this state.

I have chosen this term because it appears that many users have strong *principles-knowledge*, and hence, informational power, for acting as a sender or receiver of information but do not show the same depth of knowledge about the wider temporal, commercial, and global universe of information flow on the platform. While I am not suggesting that Twitter's new-user orientation process singularly caused this state of information flow solipsism, the one unifying aspect of Twitter that every registered user must experience does little to dispel it on the whole. With limited exception, the language that Twitter, Inc. presents to users about information flow about these wider elements is often vague, unclear, or leaves significant aspects of the flows implied or unmentioned. While the corpus gives a nod to the fact that there is a larger information ecosystem, it draws only the haziest picture of the actors that inhabit it, the flows that happen within it, the governance of that ecosystem, and the financial arrangements that shape it.

A state of information flow solipsism may have a number of serious consequences for individuals. For example, users whose principles-knowledge can be described in this way will face difficulties in exercising power in relation to the wider information ecosystem that the information they create becomes a part of; after all, it is difficult to object, protest, or consciously consent to that which you do not know about. The argument that users' *principles-knowledge* and hence, informational power, can be described as a state of information flow solipsism appears to align with Puschmann and Burgess' (2013) argument that, "End-users (both private individuals and non-profit institutions) are without a place in [the emerging data market], except in the role of passive producers of data" (p. 11). As this chapter has already listed a number of consequences for users in terms of their informational power in relation to the commercialization of user-generated content and the long-term lifecycle of information and information flows, this section will instead focus on potential outcomes for users in relation to the global ecosystem of information flow.

A state of information flow solipsism suggests that users may face difficulties in conceptualizing what information other Twitter users see as part of their own experiences of Twitter. This may be as simple as users not understanding how @replies propagate to others' timelines, or much more broadly users not understanding that, despite the statement "What you say on Twitter may be viewed all around the world instantly," that may not be true as Twitter, Inc. blocks some content in certain geographic regions. An extant state of information flow solipsism appears to undergird Lagos' (2012) argument that social media platforms such as Twitter constitute a "public sphere on steroids." Public spheres are in part based on inclusivity (at least, based on the formulation

presented by Habermas [1991]). An individual who is not fully aware that not all users on Twitter have access to the same information might think that Twitter is entirely inclusive, but this is not the case.

Without knowledge of how information does and does not flow to others, users may have difficulty putting themselves into the informational shoes of others and understanding the wider picture of differential information flow on Twitter. With a wider-base of informational power in this regard, users might protest Twitter's decision to block Tweets. In fact, many users did just this after Twitter announced it would begin regional blocking of Tweets. In response to the announcement some users began "Posting messages with the hashtags "#TwitterBlackout" and "#TwitterCensored," and "vowed to let the company know that they opposed the new policy" (Tsukayama, 2012, para. 5). Some users went so far as to organize a day of non-use to voice their displeasure. However, such fields of action were only available to individuals once they learned of the censorship.

Users in countries where the censorship takes place with the *principles-knowledge* of this facet might choose to tailor their content carefully based on a fear of being censored, or might engage ways around the censorship, such as by using a Virtual Private Network service to spoof their IP address to make it seem as though their web-traffic is originating from elsewhere. With knowledge of Twitter's regional content blocking, individuals might become more interested and involved with issues of censorship access to information around the world. But instead, these possibilities are closed off when users do not have the informational-power in regards to the broader picture of how the information they create flows.

Addressing Individual User Power

This chapter now turns to suggest a few potential solutions that could help address user informational power moving forward. Addressing users' informational power is a difficult prospect first, because not all users are the same and different users have different needs and abilities, and second, to return to Foucault, because power is relational and operates in a net-like fashion. What is meant by this is that users' power is not only dependent on its relationship with Twitter, but is also informed by other kinds of relationships, such as the relationship between Twitter and its investors, Twitter and the U.S. Federal Trade Commission, users and other social media sites, users and education institutions, etc. However, as a result of this, there are a number of avenues from which to potentially address users' informational power.

The first avenue for addressing user's informational power that stems most directly from the findings of this dissertation would be for Twitter, Inc. to internally make certain changes to its new user orientation process. Aesthetic changes, changes to the content of the policy documents, changes to the new user tutorial, and changes to the layout of the user-interface may each reduce some of the barriers to the development of users' *principles-knowledge*, and hence, informational power. The new-user registration process, and the language that appears within it, are particularly important because all registered users go through it.

However, the development of users' informational power is not something that can happen absent the user. This leads to the second avenue from which to approach the issue, changes in user behavior. Users must be willing to play an active role in the process. Here though, there appears to be an inherent problem of incentives. Given the

documented impenetrability and vagueness associated not only with Twitter's policy documents, but terms of service and privacy policies as a genre, users have few incentives to read these documents, and Twitter, Inc. has few incentives to change them (given that users are legally bound to the terms of the agreement regardless of whether or not they have read and understood it, and given the previously noted incentives that Twitter, Inc. has for not encouraging the development of *principles-knowledge* that may give users qualms). This is where the third-avenue for addressing users' informational power comes into play.

The third avenue for addressing users' informational power explores what third-party actors might do. This includes steps authorities—such as the Federal Trade Commission or legislators—might take, steps that educators engaged in digital literacy efforts might take, steps that media outlets might take, and steps that users with more developed *principles-knowledge* and hence, informational power, might take to help other users. While none of these avenues offers a silver-bullet, individually and together they offer inroads from which to challenge the status quo of users' informational power.

Changes to new-user orientation process.

There are a number of aesthetic or layout changes that Twitter, Inc. could make to the new-user orientation process that may better facilitate the development of users' *principles-knowledge* about information flows. First, adding a printable version of the Twitter Rules as one of the hyperlinked policy documents on the "Join Twitter Today" page so the entire "agreement" is actually accessible to users from that single page could be helpful to users. This would better facilitate users accessing and reading the full terms of the agreement, and hence, have a better opportunity to develop *principles-knowledge*

with fewer steps involved. Second, Twitter, Inc. might consider not using a collapsed version of the Terms of Service on the "Join Twitter Today" page, thereby forcing users to move their eyes across the length of agreement. This would at least expose users to actual length of the document and the volume of text it contains. Finally, Twitter might consider not using lighter color text or smaller fonts so that the explanations of information flows are given equal visual presence as other components.

There also changes that could be made to the content of the texts to give them greater clarity. First, several terms are used in the corpus to describe information flows on Twitter without being given a definition. For example, the terms APIs, widgets, and buttons are used in the policy documents without having ever been defined in a way that would be accessible to those who had never heard these terms before. Concrete examples of what widgets and buttons look like may be beneficial for the construction of user-knowledge of these facets. Twitter might consider spelling out some of the limitations of what API users can and cannot do with data from Twitter so that users can better mentally connect the content they create to possible outcomes for that information. Twitter might also consider spelling out its revenue generation methods in greater detail; at the very least, adding in clarification about the fact that Twitter sells access to user-generated content via the firehose APIs and that Twitter's Certified Products resell user-generated content and analysis of that content. Naming more of the partners Twitter, Inc. works with and adding links to the Certified Products program homepage could achieve greater clarity. Relatedly, Twitter, Inc. might also consider providing more specificity within its policy documents about the lifecycle of different types of information sent through Twitter. Currently, statements about information lifecycle are spread out across

the texts in a number of locations and must be pieced together, along with the Twitter's own caveats to the timelines given. The ToS and Privacy Policy may also benefit from a clearer and more in depth articulation of what information is considered public versus private, and what information is optional versus not optional. The Terms of Service might also benefit from an explicit discussion of the fact that Twitter blocks access to Tweets with a valid legal request, and the means it uses to block access. Lastly, providing greater transparency on the specific third parties Twitter works with may better allow users to build an idea of the larger information ecosystem that surrounds Twitter and the information flows on Twitter.

As Chapter 7 observed, within the Terms of Service and Privacy Policies—documents a user would encounter before they officially had an account—there are multiple instances of hyperlinks to locations that one can only access *after* one has a registered account. These hyperlinks are essentially useless to the user who is contemplating adoption; the individual who may be trying to learn about the settings *before* committing to registering. This individual is going to be put at a disadvantage in being able to find out what those account controls are and how fine grained that control will be. This problem could be solved in one of two ways: by either providing a description of the specific controls that Twitter offers within the policy documents, or (perhaps preferably) by allowing users the ability to select those settings during the sign-up process itself rather than having to wait until after they have registered.

Outside of the policy documents, there are a number of changes that could be made to the new-user tutorial that could lead users toward more robust development of *principles-knowledge* about information flows on the platform. As previously noted,

many users do not read the policy documents, so the new-user tutorial may serve as an excellent alternative or supplemental location to help users learn more about information flow on Twitter. Currently, the new-user tutorial focuses on showing new users how to tweet, how to follow, and how the timeline interface functions. It could quite easily be expanded to provide greater detail on how replies do or do not flow to followers, how trending works, how information flow is altered if a user chooses to protect their accounts, and how information flows from Twitter to third parties.

Finally, there are also changes that could be made to the structure of the interface itself. Once beyond the new-user orientation process, there is little integration of what was discussed in the policy documents with the instructional messaging on the interface. Twitter, Inc. might consider providing occasional reminders about the public nature of the platform, about the longevity of Tweets, and about the different kinds of commercial and non-commercial audiences for Tweets to facilitate the development and recall of users' *principles-knowledge*. While these proposed changes would not solve every problem of users' informational power, they may be a valuable step in combating information flow solipsism.

Users' dilemma.

Twitter's policy documents do currently provide a number of clear and specific explanations regarding information production, networking, and consumption facets of the platform. However, the corpus provides vague descriptions, implicit descriptions, highly caveated descriptions, or all together leaves out many parts of information flows on Twitter. This is not unique to Twitter however. As a genre, policy documents such as terms of service and privacy policies are often lengthy, written in "legalese," are vague,

or are otherwise difficult to read and to comprehend. In their study of web privacy policies, Jensen and Potts (2004) conclude, "only 6% of policies are readable by the most vulnerable 28.3% of the population, and that 13% of policies were only readable by people with a post-graduate education" (p. 477).

As a result of their perceived incomprehensibility, many users choose not to read the documents. As Chapter 6 made note of, 66% of registered Twitter users sampled indicated they had never read Twitter's Terms of Service and 31% indicated that they have only skimmed it. Less than 3% had read the ToS in some detail or fully. Resolving the non-reading issue is a major challenge. Even if Twitter, Inc. adopted every single one of the changes suggested in the previous section, it is hard to say how much would actually change in terms of user-knowledge, simply because many users do not read the documents. After all, there were multiple instances of gaps in users' *principles-knowledge* where the facet was explicitly mentioned within the corpus. However, it is possible that making the policy documents appear more comprehensible would increase the likelihood that users might read it. In their study of the motivations for why users read privacy policies, Milne and Culnan (2004) found, "perceived comprehension of notices had a strong effect [on whether or not users read them]" (p. 24). Hence, changes that improve the clarity of the documents may simultaneously improve users' reading rates.

While this dissertation cannot institute changes to the policy documents that lead to users reading them in more detail or create the incentives that might propel Twitter, Inc. to change to the new-user orientation process directly, third party actors may have a role to play in this regard. The next section discusses how third-party actors can play a role in incentivizing Twitter, Inc. to make a number of changes to the new-user

registration process. Further, this section will also discuss how the development of users' informational power might be addressed outside of the new-user registration process.

Third-party actors.

To return to one of the points made in Chapter 7, the context in which the texts of the new-user orientation were produced informs the particular language choices made for the pages. Twitter, Inc.'s relationship to its business investors, potential stockholders, the U.S. Federal Trade Commission, and with the norms for policy documents among Web 2.0 businesses all inform and shape the choices that Twitter, Inc. has made about its policy documents. Hence, these relationships also inform users' informational power to an extent. Changes within the relationships between Twitter, Inc. and these third-party actors may lead to different outcomes as these texts are revisited and revised.

Entities that have a measure of structural power in relation to Twitter, Inc. may be in the best position to have a direct impact on the user-Twitter relationship. As noted in Chapter 7, the FTC's enforcement authority and its Fair Information Practice Principles serve as an influential factor that has shaped the language of terms of service and privacy policies. However, the Fair Information Practice Principles are only that, principles. They are not currently enforceable by law (the FTC only has the power to take action against companies that violate their own stated policies). The generation of new law that mandates the clear articulation of information collection and use online in terms of service and privacy policies may incentivize companies like Twitter to alter their discursive practices. In 2012, U.S. President Barack Obama proposed such an effort, dubbed, the "Consumer Privacy Bill of Rights," "as part of a comprehensive blueprint to improve consumers' privacy protections and ensure that the Internet remains an engine

for innovation and economic growth" (National Telecommunications & Information Administration, 2012, para. 1). One of the goals of this "Bill of Rights" is to give consumers a better understanding of what companies that handle personal information do with that information, as well as set a number of principles regarding how businesses should and should not use consumer data. One of the foundational tenants of this bill would be that, "Consumers have a right to *easily understandable* information about privacy and security practices [emphasis added]" (Meece, 2012). By legislating a requirement on the readability of descriptions of certain information flows, this Consumer Privacy Bill of Rights may lead to better outcomes for the development of users' informational power. Unfortunately, despite fanfare from consumer groups about the 2012 proposal, the 2015 legislation put forward by the U.S. Department of Commerce based on the 2012 report has been called watered-down, unworkable, and been noted as lacking "meaningful protections for consumers" (Electronic Privacy Information Center, 2015, para. 1). While a top-down approach may help address the issue of users' informational power, it remains uncertain what the true impact of such legislation would be until it exists.

Top-down structural changes are not the only means by which informational solipsism can be challenged. As Chapter 2 made note of, users build their *principles-knowledge* via a number of different mechanisms, such as by discussing technology with friends and family, and by consuming messages about technology from media outlets or from educators, and by watching other users use a given technology. These additional avenues open up ways to promote *principles-knowledge*, and hence, informational power, outside of the context of the new-user orientation.

Digital literacy education efforts may function as an important inroad for addressing users' informational power. There has been an ongoing debate about the definition and boundaries of what exactly should constitute digital literacy and digital literacy education. Many of the proposed educational frameworks include a concept called information literacy (Eshet, 2004). It is in this area that users' *principles-knowledge* of information flows on social media platforms can fit into these existing efforts. However, just like "digital literacy," there are many alternative models for conceptualizing information literacy. The American Libraries Association, for example, defines information literacy as "a set of abilities requiring individuals to 'recognize when information is needed and have the ability to locate, evaluate, and use effectively the needed information" (American Library Association, 2000, p. 2). This approach seems to focus more narrowly on efforts that might develop *how-to knowledge* rather than *principles-knowledge*. Shapiro and Hughes (1996) critique the conceptualization of information literacy "as defined by representatives of the library community" (p. 1). Instead, the authors suggest:

> Information and computer literacy, in the conventional sense, are functionally valuable technical skills. But information literacy should in fact be conceived more broadly as a new liberal art that extends from knowing how to use computers and access information to critical reflection on the nature of information itself, its technical infrastructure, and its social, cultural and even philosophical context and impact - as essential to the mental framework of the educated information-age citizen as the trivium of basic liberal arts (grammar, logic and rhetoric) was to the educated person in medieval society. Indeed, such an extended notion of information literacy is essential to the future of democracy, if citizens are to be intelligent shapers of the information society rather than its pawns, and to humanistic culture, if information is to be part of a meaningful existence rather than a routine of production and consumption. (p. 3)

It is in this sense of information literacy as a liberal art that attention to *principles-knowledge* of information flows may be important as part of digital literacy efforts.

Educators working in areas of digital literacy may be valuable actors in raising awareness about the commodification of user-generated information, the lifecycle of information flow including long-term storage, and in addressing information flow solipsism, not just in the regards to Twitter, but potentially within Web 2.0 platforms broadly. Ultimately, such educational efforts can connect students to debates about the constitution and nature of power within the contemporary information society.

Reporters, news agencies, and mass media outlets can also play an important role in helping raise awareness about these issues. Historically, media outlets have generally not provided technical accounts of information flow on Twitter. Arceneaux and Weiss (2010) observed that news articles about the platform between 2006-2009 instead focused on aspects such as the novelty of the platform. Despite this popular media outlets can contribute to challenging the status quo of users' *principles-knowledge* and informational power by offering stories that highlight the commodification of user-generated content, that explore the long-term implications of Tweets, that describe differential information access via the platform, or address facets of the global information ecosystem that surrounds Twitter. In fact, there have been a number of such articles that have appeared since 2009 in outlets such as the *New York Times* (for example, see: Goel, 2015a, 2015b; Ronson, 2015) the *Washington Post* (for example, see: Tsukayama, 2012, 2013), and on popular web-based media such as *Fusion* (for example, see: Roose, 2015). Perhaps now that Twitter has gained greater prominence across the globe and is a bit more established, these kinds of accounts of information flows on the platform will continue to flourish.

Lastly, users with more robustly developed *principles-knowledge* may be able to help other users with less developed *principles-knowledge*. For example, the

#TwitterBlackout and #TwitterCensored user-generated awareness campaigns may have helped heighten knowledge among users about Twitter, Inc.'s implementation of policies that restrict Tweets in certain geographic areas (Tsukayama, 2012). Roose (2015) writes about a Twitter user who made the code he used to automatically delete Tweets after a certain time publicly available, thereby raising the visibility of the long-term storage of Tweets and opening a field of action for interested users to gain some measure of control over the lifecycle. Other such coordinated protest activities and tool sharing on Twitter itself may expose other users to these facets and facilitate construction of knowledge.

Future Directions for Research

As noted in the beginning of this chapter, there are several limitations of this project in terms of its generalizability. These limitations are in part due to the population used for the survey and are in part due to the limited scope of the corpus. As a result, while this study has identified a number of potential issues of concern for users' informational power, it remains unknown how widespread these issues are across the wider population of U.S. Twitter users over the age of 18. Further study with a truly random sample is needed in order to make that determination. Additionally, while the discourse analysis considers the new-user orientation process, there are many other locations within the scope of Twitter.com where Twitter, Inc. provides descriptions of information flows on the platform. For example, the "Help Center," the Twitter Blog, and Twitter's webpages for business partners all contain discussions of information flows that may potentially impact the construction of users' *principles-knowledge* (though users are not as universally exposed to these areas as they are to the new user registration process). Further analysis and reflection on these other texts may also yield insights into how

Twitter, Inc. talks about information flows in different situational contexts. Relatedly, as the mobile version of Twitter has been gaining prominence in use, particularly since Twitter, Inc. introduced its own official mobile app, the text from this interface may also be a fruitful space to explore how information flows on Twitter are depicted.

Another limitation of this project is that it ties the arguments about users' informational power to hypothetical action that users might take. One future direction for this work is to build on the findings in this study by exploring how different states of user knowledge correlate with different use behaviors. The survey data does contain some information about different kinds of use and self-categorization of use by respondents that will be explored in greater detail in future analysis. However, parallel experimental, observational, or "trace ethnographic" study (Geiger & Ribes, 2011) may better help articulate the relationship between *principles-knowledge*, informational power, and fields of action that are actually pursued by users, not just ones that are possible.

This work may also raise a number of questions about what "user intent" looks like in the contemporary social media environment. In 2010, Twitter, Inc. announced that it was giving the archive of all Tweets ever generated to the Library of Congress. In this announcement, Biz Stone, one of the founders of Twitter stated: "… most of these tweets are created with the intent that they will be publicly available" (Stone, 2010, para. 2). However, as was seen in Chapter 6, nearly one half of the registered Twitter users surveyed did not know whether Twitter "protects" Tweets by default. With nearly half of registered users being unsure or incorrect in their understanding of whether Tweets are "protected" by default, and more than 60% still unaware of the fact that the Library of Congress archives Tweets four years after the agreement was announced, the assertion of

intent seems problematic. This work may also have import to discussions about users' "informed consent," about the ethics of Twitter data use in scholarly research, and about the use of the Library of Congress's Twitter archive.

This project has provided a snapshot of some of the issues around users' understanding of the Twitter platform, Twitter's own discourse about its product, and about the information flows of the platform. However, further research into other new media platforms is needed; specifically, research that explores how public understanding of information flow varies across social media platforms in relation to different organizational discourse practices and different site structures. For example, the Reddit and Tumblr platforms have been growing in use and importance in the social media landscape, yet remain understudied in comparison to spaces like Facebook and Twitter. The companies that purvey these technologies also have very different approaches to engaging users, in the language they present to users, in their transparency, and in their structuring of information flows on the sites. Another possible direction this line of research could be taken in is to explore whether the different language use patterns and interface designs relate to different kinds of user knowledge about information flows on these platforms, or if many of the same trends and findings from this study of Twitter users emerge in these spaces.

In terms of the theoretical framework of this dissertation, this work could be expanded further through the integration of Braman's (2006) work on the different phases of power. In *Change of State*, Braman distinguishes between three phases of power: actual, "power that is currently being exercised" (p. 28); potential, "claimed resources and techniques of power that are not currently in use" (p. 28); and virtual, "resources and

techniques of power that are not currently extant but that might be brought into existence using available resources and knowledge" (p. 28). Such a framework could be highly useful in investigating the link between user-knowledge and user-behavior in greater depth.

Lastly, the diffusion approach of this study could also be expanded or contrasted with alternative approaches to understanding the relationship between social actors, meaning, and technology from within STS. Approaches such as "technological frames" (Bijker, 1995, 2001; Orlikowski & Gash, 1994) may yield insights into how Twitter gains an overall social "meaning" as a technology, which begins to fix the obduracy of the platform. The strength of this framework is that it also considers elements such as users' practice in the social constitution of the technological artifact—an element not given considerable attention in this dissertation.

Conclusions

This dissertation has addressed three operationalized research questions: What knowledge of information flows do users have? How does the technological discourse surrounding the site created by Twitter's business purveyors describe information flows? And finally, in the user-Twitter relationship, what is the state of Twitter users' informational power in regards to the informational flows of the platform? In answering these three questions, this work has unpacked power in the user-Twitter relationship by exploring how information flows on Twitter, users' *principles-knowledge* of information flows, how Twitter's new user orientation depicts information flows, and what the some of the potential consequences of the junctures and disjunctures between these elements are for users' informational power and available fields of action. This research suggests

that in the user-Twitter relationship, users have strong *principles-knowledge* about information production, consumption, and networking on the platform, but have far weaker bases of *principles-knowledge* regarding how information is made available or appears to other Twitter users, to third-parties, becomes commodified by Twitter, Inc. and by Twitter, Inc.'s partners, and the long-term lifecycle of user-generated content. The new-user orientation, on the whole, offers little to counter these knowledge gaps. As a result, users appear to have strong informational power in regards to their own production, consumption, and networking, but relatively weaker informational power in relation to the wider commodification of content, to the long-term of the information they create, and to the larger information ecosystem of and surrounding Twitter.

This dissertation lays out a number of potential consequences registered users may face as a result. While users may be well positioned to do things like take advantage of Twitter's affordances in order to exercise symbolic power through the medium, they have a far less developed base of informational power to potentially challenge the commercialization of the content they generate, to be able to evaluate the potential long-term implications of their information content production, and to understand and challenge the restrictions on the global flow of information on Twitter. Information flow solipsism is a pressing issue because, if social media sites such as Twitter become further entrenched as dominant vehicles for communication, our knowledge of how these technologies function will play an ever increasing role in our abilities to make purposeful and meaningful choices about the use and governance of these spaces.

The first chapter of this dissertation began with a story about a controversy among a handful of users when they discovered the Library of Congress would be archiving

Twitter. Their expressions of displeasure could only take place because these users learned that Twitter was giving the LoC all public Tweets ever created. In many ways, controversies such as this serve an important function in raising awareness of how information flows and in users then making changes to (or choosing not to change) their behaviors on the basis of that new informational power. However, relying on moments of controversy or crisis to serve as educational tools seems a cruel method for furthering users' knowledge. In many cases, these situations only arise because a "bad" outcome has already occurred and the controversy now reveals lessons to be learned from. For example, it was not until after the controversy around whether Twitter was actively censoring the #Occupy related trending topics that some protest organizers began thinking more strategically about messaging on the platform. However, for Occupy organizers, such insight was likely too little, too late. What this dissertation offers is a mapping of users' informational power that may be helpful in understanding some of the roots of these events, and may be useful in tracing where there may be extant problems in users' informational power before the unwelcome outcome. It also offers a number of suggestions that may be helpful in combatting the more egregious gaps in users' *principles-knowledge* and hence, informational power.

The first chapter of this dissertation also introduces J. Cohen's (2012) argument that "the emerging regime of information rights and privileges … allows individuals less and less control over information flows to, from, and about themselves" (p. 3). The findings from this analysis suggest that diminished control over information flows to, from, and about one's self may not just be realized through legal regimes, but in the case of Twitter, is also realized through the shaping of individuals' *principles-knowledge* of

information flows, which thereby shapes the individual's informational power. Without a robustly developed base of knowledge regarding the different facets of information flows—knowing how one's data appears to others, is shared or sold, stored and archived—it becomes much more difficult for an individual to judge the potential consequences of engaging in the use of Twitter, and much more difficult for an individual to enter fields of action in which they exert measures of power, control, or influence over information flows in relation to the sociotechnical platform.

While some have praised the Twitter platform for its relative simplicity and transparency in comparison to other social media sites, this dissertation has shown that this does not necessarily mean that users have a robustly developed set of *principles-knowledge* about information flows on the platform. This work demonstrates that this so-called simplicity and transparency includes descriptions of information flows that are frequently vague, imprecise, or leave significant elements such as the commercialization of user-generated content, the long-term of Tweets, and the wider picture of information flows beyond the user unsaid. This so called simplicity and transparency has furthered a state of informational power among users in the user-Twitter relationship perhaps best described as "information flow solipsism." And if, indeed, Twitter is a shining beacon of simplicity and transparency among Web 2.0 sites, there is much to be concerned about for users' power in the contemporary social media landscape.

References

5th Criminal Court of Ankara. (2014). Turkish notice regarding two Twitter accounts. Chilling Effects. Retrieved August 21, 2014, from https://www.chillingeffects.org/international/notice.cgi?NoticeID=1971282

Achrekar, H., Gandhe, A., Lazarus, R., Yu, S.-H., & Liu, B. (2011). Predicting flu trends using Twitter data. Presented at the Computer Communications Workshops (Information Workshops), 2011 IEEE Conference, (pp. 702–707) IEEE.

Acquisti, A., & Gross, R. (2006). Imagined communities: Awareness, information sharing, and privacy on the Facebook. *Proceedings of the 6th Workshop on Privacy Enhancing Technologies*, *4258*, 36–58.

Adams, R. C. (1989). *Social survey methods for mass media research*. Mahwah, NJ: Lawrence Erlbaum.

Albrechtslund, A. (2008). Online social networking as participatory surveillance. *First Monday*, *13(3)*. Retrieved from http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2142/1949

Alexa.com. (2015). Alexa Top 500 global sites. Retrieved February 28, 2015, from http://www.alexa.com/topsites

Allen, E. (2013, January 4). Update on the Twitter archive at the Library of Congress | Library of Congress Blog [webpage]. Retrieved April 3, 2013, from http://blogs.loc.gov/loc/2013/01/update-on-the-twitter-archive-at-the-library-of-congress/?doing_wp_cron=1365013571.0066039562225341796875

American Library Association, A. L. (2000). *Information literacy competency standards for higher education*. Chicago, IL.

Andersen, P. (2007). What is Web 2.0?: Ideas, technologies and implications for education. *JISC Technology and Standards Watch*. Retrieved from http://www.jisc.ac.uk/media/documents/techwatch/tsw0701b.pdf

Andrejevic, M. (2013). Alienation's returns. In C. Fuchs & M. Sandoval (Eds.), *Critique, Social Media and the Information Society* (pp. 179 – 190). London, England: Routledge.

Ante, S. E., & Das, A. E. A. A. (2011, February 10). Twitter as tech bubble barometer. *Wall Street Journal*. Retrieved from http://online.wsj.com/news/articles/SB10001424052748703716904576134543029279426?KEYWORDS=twitter&mg=reno64-wsj&url=http%3A%2F%2Fonline.wsj.com%2Farticle%2FSB10001424052748703716904576134543029279426.html%3FKEYWORDS%3Dtwitter

Arceneaux, N., & Weiss, A. S. (2010). Seems stupid until you try it: Press coverage of Twitter, 2006-9. *New Media and Society*, *12*(8), 1262–1279.

Armstrong, C. L., & Gao, F. (2010). Now Tweet This How News Organizations Use Twitter. *Electronic News*, *4*(4), 218–235.

Asur, S., Huberman, B. A., Szabo, G., & Wang, C. (2011). Trends in social media: Persistence and decay. In *5th International AAAI Conference on Weblogs and Social Media*.

Bachrach, P., & Baratz, M. S. (1962). Two faces of power. *American Political Science Review*, *56*(04), 947–952.

Bajarin, B. (2013). The value of Twitter. Retrieved April 27, 2014, from http://techland.time.com/2013/05/28/the-value-of-twitter/

Baker, A. E. (2011). Ethical considerations in Web 2.0 archives. *SLIS Student Research Journal*, *1*(1), 4.

Bakshi, H. (2011). *Framework for crawling and local event detection using Twitter data* (M.S., Graduate School - New Brunswick). Rutgers The State University of New Jersey - New Brunswick, United States -- New Jersey. Retrieved from http://search.proquest.com.ezproxy.lib.uwm.edu/pqdtft/docview/897902775/abstract/1351330AA0A7C13B6E2/1?accountid=15078

Barlow, J. P. (1994). A taxonomy of information. *Bulletin of the American Society for Information Science*, *20*(5), 13–17.

Barnlund, D. C. (1970). A transactional model of communication. *Foundations of Communication Theory*, 83–102.

Bates, M. (2005). Information and knowledge: An evolutionary framework for information science. *Information Research*, *10*(4), 10–4.

Bates, M. (2006). Fundamental forms of information. *Journal of the American Society for Information Science & Technology.*, *57*(8), 1033–1045.

Bazerman, C. (1998). The production of technology and the production of human meaning. *Journal of Business and Technical Communication*, *12*, 381–387.

Beevolve. (2012). An exhaustive study of Twitter users across the world. *Social Media Analytics - Beevolve*. Retrieved April 11, 2014, from http://www.beevolve.com/twitter-statistics/#d2

Bendix, R. (1978). *Max Weber: An intellectual portrait*. Berkeley, CA: University of California Press.

Benkler, Y. (2007). *The wealth of networks: How social production transforms markets and freedom*. New Haven, CT: Yale University Press.

Bennett, S. (2012). 55% of journalists worldwide use Twitter, Facebook to source news stories. *AllTwitter*. Retrieved April 27, 2014, from http://www.mediabistro.com/alltwitter/digital-journalism_b24440

Bercovici, J. (2014). Dick Costolo's four-part plan for making "A Better Twitter." *Forbes*. Retrieved April 27, 2014, from http://www.forbes.com/sites/jeffbercovici/2014/02/05/dick-costolos-four-part-plan-for-making-a-better-twitter/

Berlo, D. K. (1960). *The process of communication: an introduction to theory and practice*. New York: Holt, Rinehart and Winston.

Bijker, W. E. (1995). *Of bicycles, bakelites, and bulbs: Toward a theory of sociotechnical change*. Cambridge, MA: MIT Press.

Bijker, W. E. (2001). Understanding technological culture through a constructivist view of science, technology, and society. *Visions of STS: Counterpoints in Science, Technology and Society Studies*.

Bilton, N. (2012). Disruptions: Facebook users ask, "Where's our cut?". *New York Times Bits Blog*. Retrieved from http://bits.blogs.nytimes.com/2012/02/05/disruptions-facebook-users-ask-wheres-our-cut/

Bilton, N. (2013, October 9). All is fair in love and Twitter. *The New York Times*. Retrieved from http://www.nytimes.com/2013/10/13/magazine/all-is-fair-in-love-and-twitter.html

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, *2*(1), 1–8. http://doi.org/10.1016/j.jocs.2010.12.007

Boudreaux, D. J., & Crampton, E. (2003). Truth and consequences: Some economics of false consciousness. *INDEPENDENT REVIEW-OAKLAND-*, *8*(1), 27–46.

boyd, danah. (2009). Twitter: "pointless babble" or peripheral awareness + social grooming? Retrieved May 11, 2014, from http://www.zephoria.org/thoughts/archives/2009/08/16/twitter_pointle.html

boyd, danah, & Crawford, K. (2011). Six provocations for big data. Presented at the A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society, Oxford Internet Institute.

boyd, danah, Golder, S., & Lotan, G. (2010). Tweet Tweet Retweet: Conversational aspects of retweeting on Twitter. In *Proceedings of HICSS-43*. Kauai, HI. Retrieved from http://www.danah.org/papers/TweetTweetRetweet.pdf

boyd, danah, & Hargittai, E. (2010). Facebook privacy settings: Who cares? *First Monday*, *15*(8). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/3086

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, *15*(5), 662–679.

Braman, S. (1989). Defining information: An approach for policymakers. *Telecommunications Policy*, *13*(3), 233–242.

Braman, S. (2006). *Change of state: Information, policy, and power*. Cambridge, MA: MIT Press.

Brenner, J., & Smith, A. (2013). 72% of online adults are social networking site users. Retrieved August 25, 2013, from http://pewinternet.org/Reports/2013/social-networking-sites.aspx

Bruns, A. (2008). *Blogs, Wikipedia, Second Life, and beyond: From production to produsage* (Vol. 45). New York, NY: Peter Lang.

Bruns, A., Burgess, J. E., Crawford, K., & Shaw, F. (2012). *# qldfloods and@ QPSMedia: Crisis communication on Twitter in the 2011 south east Queensland floods*. ARC Centre of Excellence for Creative Industries and Innovation. Retrieved from http://eprints.qut.edu.au/48241

Bucher, T. (2012). *Programmed sociality: A software studies perspective on social networking sites*. University of Oslo.

Buckland, M. K. (1991). Information as thing. *Journal of the American Society for Information Science*, *42*(5), 351–360.

Burns, A., & Eltham, B. (2009). Twitter free Iran: An evaluation of Twitter's role in public diplomacy and information operations in Iran's 2009 Election Crisis. Presented at the Communications Policy & Research Forum, Sydney, Australia. Retrieved from http://eprints.vu.edu.au/15230/

Cameron, D., & Kulick, D. (2003). *Language and sexuality*. Cambridge, UK: Cambridge University Press.

Castells, M. (2009). *Communication power*. Cary, NC: Oxford University Press, USA.

Chan, C. (2012). It's not a crime to break a terms of service agreement (so it's okay to never read them). Retrieved January 6, 2015, from

http://gizmodo.com/5901339/its-not-a-crime-to-break-a-terms-of-service-agreement-so-keep-on-not-reading-them

Chandler, D. (2011). Transmission model of communication. Retrieved January 25, 2014, from http://www.aber.ac.uk/media/Documents/short/trans.html

Cheney, G., Christensen, L. T., Conrad, C., & Lair, D. (2004). The nature of rhetoric in an organizational world. In D. Grant, C. Hardy, C. Oswick, & L. L. Putnam (Eds.), *The SAGE Handbook of Organizational Discourse*. Thousand Oaks, CA: SAGE.

Cheney, G., & McMillan, J. J. (1990). Organizational rhetoric and the practice of criticism. *Journal of Applied Communication Research*, *18*(2), 93–114.

Chi, F., & Yang, N. (2011). Twitter adoption in congress. *Review of Network Economics*, *10*(1), 3.

Cohen, J. E. (2012). *Configuring the networked self: Law, code, and the play of everyday practice*. New Haven, CT: Yale University Press.

Cohen, N. S. (2008). The valorization of surveillance: Towards a political economy of Facebook. *Democratic Communiqué*, *22*(1), 5–22.

Comor, E. (2011). Contextualizing and critiquing the fantastic prosumer: Power, alienation and hegemony. *Critical Sociology*, *37*(3), 309–327. http://doi.org/10.1177/0896920510378767

Constine, J. (2012). Not at any price: Twitter denied data to Google and bet on itself. Retrieved September 16, 2013, from http://techcrunch.com/2012/01/13/twitter-google-firehose/

Cook, T. (2014, October 30). Tim Cook speaks up. *BusinessWeek: Technology*. Retrieved from http://www.businessweek.com/articles/2014-10-30/tim-cook-im-proud-to-be-gay

Coté, M., & Pybus, J. (2007). Learning to immaterial labour 2.0: MySpace and social networks. *Ephemera: Theory and Politics in Organization*, *7*(1), 88–106.

Cowan, R. S. (1987). The consumption junction: A proposal for research strategies in the sociology of technology. *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*, 261–80.

Craik, K. (1943). *The nature of explanation.* Cambridge, England: Cambridge University Press.

Cranor, L. F. (2003). P3P: Making privacy policies more useful. *IEEE Security & Privacy*, *1*(6), 50–55.

Crum, R. (2014). Twitter's 23 million "bot" question mark. Retrieved January 16, 2015, from http://blogs.marketwatch.com/thetell/2014/08/12/twitters-23-million-bot-question-mark/

Cukier, W., Ngwenyama, O., Bauer, R., & Middleton, C. (2009). A critical analysis of media discourse on information technology: preliminary results of a proposed method for critical discourse analysis. *Information Systems Journal*, *19*(2), 175–196.

Culnan, M. J., McHugh, P. J., & Zubillaga, J. I. (2010). How large US companies can use Twitter and other social media to gain business value. *MIS Quarterly Executive*, *9*(4).

Curtis, D. (2012). Twitter is tracking you on the web. Retrieved April 4, 2014, from http://dcurt.is/twitter-is-tracking-you-on-the-web

Dahl, R. A. (1957). The concept of power. *Behavioral Science*, *2*(3), 201–215.

Dash, A. (2013). What it's like being verified on Twitter. Retrieved April 19, 2014, from http://dashes.com/anil/2013/03/what-its-like-being-verified-on-twitter.html

Dreyfus, H. L., & Rabinow, P. (1982). *Michel Foucault: Beyond structuralism and hermeneutics* (Vol. 2). University of Chicago Press Chicago.

Duggan, M., Ellison, N. B., Lampe, C., Lenhart, A., & Madden, M. (2015). Social media update 2014. Retrieved from http://www.pewinternet.org/2015/01/09/social-media-update-2014/

Dybwad, B. (2009). Twitter drops "What are You Doing?" now asks "What's Happening?." Retrieved November 11, 2011, from http://mashable.com/2009/11/19/twitter-whats-happening/

Earle, P. S., Bowden, D. C., & Guy, M. (2011). Twitter earthquake detection: earthquake monitoring in a social world. *Annals of Geophysics*, *54*(6), 708–715. http://doi.org/10.4401/ag-5364

Ebner, M., Beham, G., Costa, C., & Reinhardt, W. (2009). How people are using Twitter during conferences. In *Proceedings of 5. Edumedia Conference* (pp. 145–156).

Electronic Frontier Foundation. (n.d.). Section 230 of the Communications Decency Act. Retrieved May 3, 2014, from https://www.eff.org/issues/cda230

Electronic Privacy Information Center. (2015). White House: Consumer Privacy Bill of Rights. Retrieved from https://epic.org/privacy/white_house_consumer_privacy_.html

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook "friends:" Social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication*, *12*(4), 1143–1168.

Elman, J. (2010). The power of suggestions. Retrieved January 8, 2015, from https://blog.twitter.com/2010/power-suggestions

eMarketer. (2013). Most digital ad growth now goes to mobile as desktop growth falters. *eMarketer*. Retrieved January 30, 2015, from http://www.emarketer.com/Article/Most-Digital-Ad-Growth-Now-Goes-Mobile-Desktop-Growth-Falters/1010458

Eshet, Y. (2004). Digital literacy: A conceptual framework for survival skills in the digital era. *Journal of Educational Multimedia and Hypermedia*, *13*(1), 93–106.

Fairclough, N. (1995). *Critical discourse analysis: The critical study of language*. London ; New York: Longman Pub Group.

Fairclough, N. (1999). Global capitalism and critical awareness of language. In A. Jaworski & N. Coupland (Eds.), *The Discourse Reader* (2nd Edition, pp. 146–157). New York, NY: Routledge.

Fairclough, N., & Wodak, R. (1997). Critical discourse analysis. In T. A. van Dijk (Ed.), *Discourse Studies. A multidisciplinary introduction* (Vol. 2, pp. 258–284). London: Sage.

Federal Trade Commission. (2010). Twitter settles charges that it failed to protect consumers' personal information; Company will establish independently audited information security program. Retrieved January 6, 2015, from

http://www.ftc.gov/news-events/press-releases/2010/06/twitter-settles-charges-it-failed-protect-consumers-personal

Federal Trade Commission. (2011). Docket No. C-4316: United States of American Federal Trade Commission, In the Matter of Twitter, Inc., a corporation. Retrieved from http://www.ftc.gov/sites/default/files/documents/cases/2011/03/110311twittercmpt.pdf

Feenberg, A. (1992). Subversive rationalization: Technology, power, and democracy. *Inquiry*, *35*(3-4), 301–322. http://doi.org/10.1080/00201749208602296

Fernback, J., & Papacharissi, Z. (2007). Online privacy as legal safeguard: the relationship among consumer, online portal, and privacy policies. *New Media & Society*, *9*(5), 715–734.

Fiske, J. (1982). Communication theory. In *Introduction to Communication Studies, John Fiske, Ed. London/New York, Methuen*, pp. 6–24.

Fiske, J. (1987). *Television culture*. New York, NY: Methuen.

Fiveash, K. (2015). Twitter complies with Turkey's "national security" blackout demand – BLOCKS newspaper's tweets. *The Register.* Retrieved January 19, 2015, from http://www.theregister.co.uk/2015/01/18/twitter_complies_with_turkey_censorship_demands/

Foucault, M. (1972). *The archaeology of knowledge*. New York, NY: Pantheon.

Foucault, M. (1973). *The Birth of the Clinic, trans*. New York, NY: Pantheon.

Foucault, M. (1977). *Discipline and punish: The birth of the prison*. New York, NY: Vintage Books.

Foucault, M. (1978). *The history of sexuality, Vol. 1*. New York, NY: Random House.

Foucault, M. (1979). Power and norm: notes. *Michel Foucault: Power, Truth, Strategy. Sydney: Feral*, 59–66.

Foucault, M. (1980). *Power/Knowledge: Selected interviews and other writings, 1972-1977*. (C. Gordon, Ed.) (First American Edition). New York, NY: Vintage.

Foucault, M. (1988). *Madness and civilization: A history of insanity in the age of reason*. New York, NY: Random House.

Fox, B. I., & Varadarajan, R. (2011). Use of Twitter to encourage interaction in a multi-campus pharmacy management course. *American Journal of Pharmaceutical Education*, *75*(5).

Fox, S. (2000). Trust and privacy online. Retrieved October 9, 2010, from http://www.pewinternet.org/Reports/2000/Trust-and-Privacy-Online.aspx

French, J. R. P., & Raven, B. (1959). The bases of social power. In *Studies in social power* (pp. 150–167). Ann Arbor. MI: University of Michigan Press.

Frommer, D., & Angelova, K. A. M. (2009). Half of YouTube videos get fewer than 500 views. Retrieved July 19, 2014, from http://www.businessinsider.com/chart-of-the-day-youtube-videos-by-views-2009-5

Fuchs, C. (2008). *Internet and society: Social theory in the information age* (Reprint). New York, NY: Routledge.

Fuchs, C. (2009a). Social networking sites and the surveillance society. *A Critical Case Study of the Usage of studiVZ, Facebook, and MySpace by Students in Salzburg in the Context of Electronic Surveillance. Salzburg/Vienna: Research Group UTI*.

Fuchs, C. (2010a). Social software and Web 2.0: Their sociological foundations and implications. *Handbook of Research on Web, 2*(3.0), 764–789.

Fuchs, C. (2010b). Web 2.0, prosumption, and surveillance. *Surveillance & Society*, *8*(3), 288–309.

Fuchs, C. (2014). *Social media: A critical introduction*. New York, NY: Sage.

Gadde, V. (2014). Challenging the access ban in Turkey. Retrieved May 3, 2014, from https://blog.twitter.com/2014/challenging-the-access-ban-in-turkey

Gaffney, D., & Puschmann, C. (2013). Data collection on Twitter. In K. Weller, A. Bruns, J. Burgess, & M. Mahrt (Eds.), *Twitter and Society*. New York, NY: Peter Lang.

Galloway, A. R. (2006). *Protocol: How control exists after decentralization*. Boston, MA: The MIT Press.

Gaver, W. W. (1991). Technology affordances. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 79–84). ACM.

Gee, J. P. (2010a). *An introduction to discourse analysis: Theory and method* (3rd edition). Milton Park, Abingdon: Routledge.

Gee, J. P. (2010b). *How to do discourse analysis: A toolkit*. New York, NY: Routledge.

Geiger, R. S., & Ribes, D. (2011). Trace ethnography: Following coordination through documentary practices. In *System Sciences (HICSS), 2011 44th Hawaii International Conference on* (pp. 1–10). IEEE.

Gibson, J. J. (1977). The concept of affordances. *Perceiving, Acting, and Knowing*, 67–82.

Gillespie, T. (2010). The politics of "platforms." *New Media & Society*, *12*(3), 347–364.

Gill, R. (2000). Discourse analysis. In M. Bauer & G. Gaskell (Eds.), *Qualitative Researching With Text, Image and Sound* (pp. 172–190). Thousand Oaks, CA: Sage Publications Ltd.

Goel, V. (2015a). Twitter's chief sees plenty of money in Tweet flow [The New York Times]. Retrieved from http://bits.blogs.nytimes.com/2015/02/11/twitters-chief-sees-plenty-of-money-in-tweet-flow/

Goel, V. (2015b). Twitter traffic reflects what people like on TV, study says [The New York Times]. Retrieved from http://bits.blogs.nytimes.com/2015/03/08/twitter-chatter-reveals-what-we-like-to-watch/

Golbeck, J., Grimes, J. M., & Rogers, A. (2010). Twitter use by the US Congress. *Journal of the American Society for Information Science and Technology*, *61*(8), 1612–1621.

González, V. M., & Juárez, R. (2013). How do you understand Twitter?: Analyzing mental models, understanding and learning about complex interactive systems. In *Human Computer Interaction* (pp. 103–110). Springer.

Gramsci, A. (1995). *Further selections from the prison notebooks*. Minneapolis, MN: University of Minnesota Press.

Grant, W. J., Moon, B., & Busby Grant, J. (2010). Digital dialogue? Australian politicians' use of the social network tool Twitter. *Australian Journal of Political Science*, *45*(4), 579–604.

Grosseck, G., & Holotescu, C. (2008). Can we use Twitter for educational activities. In *4th international scientific conference, eLearning and software for education, Bucharest, Romania*.

Grossman, L. (2009). Iran protests: Twitter, the medium of the movement. *Time Magazine*, *17*.

Habermas, J. (1991). *The Structural transformation of the public sphere: An inquiry into a category of bourgeois society*. Cambridge, MA: MIT Press.

Habermas, J. (2006). Political communication in media society: Does democracy still enjoy an epistemic dimension? The impact of normative theory on empirical research. *Communication Theory*, *16*(4), 411–426.

Hall, S. (2001). Foucault: Power, knowledge and discourse. In M. Wetherell, S. Taylor, & S. J. Yates (Eds.), *Discourse Theory and Practice: A Reader*. SAGE.

Hambrick, M. E., Simmons, J. M., Greenhalgh, G. P., & Greenwell, T. C. (2010). Understanding professional athletes' use of Twitter: A content analysis of athlete Tweets. *International Journal of Sport Communication*, *3*(4), 454–471.

Hansen, L. (2010). The Library Of Congress is archiving your Tweets. *NPR*. Retrieved May 22, 2013, from http://www.npr.org/templates/story/story.php?storyId=126086325

Hardt, M., & Negri, A. (2005). *Multitude: War and democracy in the age of empire*. New York, NY: Penguin (Non-Classics).

Hargittai, E. (2009). An update on survey measures of web-oriented digital literacy. *Social Science Computer Review*, *27*(1), 130–137.

Hargittai, E., & Litt, E. (2011). The tweet smell of celebrity success: Explaining variation in Twitter adoption among a diverse group of young adults. *New Media & Society*, *13*(5), 824–842. http://doi.org/10.1177/1461444811405805

Harkinson, J. (2013). Here's how Twitter can track you on all of your devices. *Mother Jones*. Retrieved April 6, 2014, from http://www.motherjones.com/politics/2013/09/twitter-could-threaten-your-privacy-more-facebook

Harrington, S., Highfield, T., & Bruns, A. (2012). More than a backchannel: Twitter and television. *Ed. José Manuel Noguera. Audience Interactivity and Participation. COST Action ISO906 Transforming Audiences, Transforming Societies, Brussels, Belgium*, 13–17.

Haugaard, M. (2002). *Power: A reader* (annotated edition). Manchester, England: Manchester University Press.

Hayles, N. K. (2010). *My mother was a computer: Digital subjects and literary texts*. Chicago, IL: University of Chicago Press.

Highfield, T., Harrington, S., & Bruns, A. (2013). Twitter as a technology for audiencing and fandom: The #Eurovision phenomenon. *Information, Communication & Society*, *16*(3), 315–339.

Himelboim, I. (2014, February 20). Mapping Twitter topic networks: From polarized crowds to community clusters. *Pew Internet and American Life Project*. Retrieved from http://www.pewinternet.org/2014/02/20/mapping-twitter-topic-networks-from-polarized-crowds-to-community-clusters/

Holmes, R. (2013). Selling out or growing up? Why paid social media is set to surge. Retrieved September 16, 2013, from http://www.forbes.com/sites/ciocentral/2013/02/25/selling-out-or-growing-up-why-paid-social-media-is-set-to-surge/

Hong, S., & Nadler, D. (2011). Does the early bird move the polls?: The use of the social media tool 'Twitter' by US politicians and its impact on public opinion. In *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times* (pp. 182–186). ACM.

Horaczek, S. (2012). People upload an average of 250 million photos per day to Facebook. Retrieved August 30, 2013, from http://www.popphoto.com/news/2012/02/people-upload-average-250-million-photos-day-to-facebook

Huffington Post. (2010). Twitter user statistics revealed. *Huffington Post.* Retrieved August 5, 2013, from http://www.huffingtonpost.com/2010/04/14/twitter-user-statistics-r_n_537992.html

Hughes, T. P. (1986). The seamless web: Technology, science, etcetera, etcetera. *Social Studies of Science*, *16*(2), 281–292. http://doi.org/10.1177/0306312786016002004

Hughes, T. P. (1987). The evolution of large technological systems. *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*, 51–82.

Hull, G., Lipford, H. R., & Latulipe, C. (2011). Contextual gaps: privacy issues on Facebook. *Ethics and Information Technology*, *13*(4), 289–302.

Ifukor, P. (2010). "Elections" or "Selections"? Blogging and Twittering the Nigerian 2007 general elections. *Bulletin of Science, Technology & Society*, *30*(6), 398 – 414. http://doi.org/10.1177/0270467610380008

Isaac, M. (2013a). Dick Costolo Wrote the Twitter Playbook — Two Years Before He Joined the Company. *AllThingsD.* Retrieved from http://allthingsd.com/20130913/dick-costolo-wrote-the-twitter-playbook-two-years-before-he-joined-the-company/

Isaac, M. (2013b). Twitter Tweaks Its Welcome Screen to Better Explain Itself. *AllThingsD.* Retrieved from http://allthingsd.com/20131007/twitter-tweaks-its-welcome-screen-to-better-explain-itself/

IT Law Group. (2014). Lessons from FTC v. Twitter. *IT Law Group*. Retrieved January 6, 2015, from http://www.itlawgroup.com/resources/articles/99-lessons-from-ftc-v-twitter

Jackson, N., & Lilleker, D. (2011). Microblogging, constituency service and impression management: UK MPs and the use of Twitter. *The Journal of Legislative Studies, 17*(1), 86–105.

Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology*, *60*(11), 2169–2188.

Jarrett, K. (2008). Interactivity is Evil! A critical investigation of Web 2.0. *First Monday*, *13*(3). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/2140

Java, A., Song, X., Finin, T., & Tseng, B. (2007). Why we Twitter: Understanding microblogging usage and communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis* (pp. 56–65).

Jaworski, A., & Coupland, N. (1999). *The discourse reader* (Vol. 2). London, England: Routledge London.

Jenkins, H., & Deuze, M. (2008). Convergence culture. *Convergence: The International Journal of Research into New Media Technologies*, *14*(1), 5–12.

Jensen, C., & Potts, C. (2004). Privacy policies as decision-making tools: An evaluation of online privacy notices. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 471–478). ACM.

Johnstone, B. (2008). *Discourse analysis*. Malden, MA: Blackwell.

Junco, R., Elavsky, C. M., & Heiberger, G. (2013). Putting twitter to the test: Assessing outcomes for student collaboration, engagement and success. *British Journal of Educational Technology*, *44*(2), 273–287. http://doi.org/10.1111/j.1467-8535.2012.01284.x

Juris, J. S. (2012). Reflections on# Occupy Everywhere: Social media, public space, and emerging logics of aggregation. *American Ethnologist*, *39*(2), 259–279.

Kafka, P. (2013). Twitter hikes its promoted trend prices again, to $200,000 a day. *AllThingsD.* Retrieved September 16, 2013, from http://allthingsd.com/20130209/twitter-hikes-its-promoted-trend-prices-again-to-200000-a-day/

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, *53*(1), 59–68. http://doi.org/10.1016/j.bushor.2009.09.003

Kim, S. (2013, October 4). Twitter IPO filing, will trade under TWTR. *ABCNews*. Retrieved May 2, 2014, from http://abcnews.go.com/Business/twitter-ipo-filing-reveals-500-million-tweets-day/story?id=20460493

Kline, S. (2003). What is technology? In R. Scharff & V. Dusek (Eds.), *Philosophy of Technology: The Technological Condition - An Anthology* (pp. 210–212). Wiley-Blackwell.

Koh, Y. (2014, April 11). Report: 44% of Twitter accounts have never sent a Tweet. *Wall Street Journal*. Retrieved from http://blogs.wsj.com/digits/2014/04/11/new-data-quantifies-dearth-of-tweeters-on-twitter/?mod=e2tw&utm_source=twitterfeed&utm_medium=twitter

Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media? In *The 19th World-Wide Web (WWW) Conference*. Raleigh, North Carolina. Retrieved from http://an.kaist.ac.kr/ haewoon/papers/2010-www-twitter.pdf

Kwon, E. S., & Sung, Y. (2011). Follow me! Global marketers' Twitter use. *Journal of Interactive Advertising, 12*(1), 4–16.

Lagos, T. (2012). Social media — the public sphere on steroids. *Seattle Times*. Retrieved February 17, 2015, from http://seattletimes.com/html/opinion/2018143386_guest06lagos.html

Lampos, V., De Bie, T., & Cristianini, N. (2010). Flu detector-tracking epidemics on Twitter. *Machine Learning and Knowledge Discovery in Databases*, 599–602.

Langlois, G., McKelvey, F., Elmer, G., & Werbin, K. (2009). Mapping commercial Web 2.0 worlds: Towards a new critical ontogenesis. *Fibreculture*, *14*, 1–14.

Latour, B. (2007). *Reassembling the social: An introduction to actor-network-theory*. Oxford, England: Oxford University Press.

Lessig, L. (2006). *Code: And other laws of cyberspace, version 2.0*. New York, NY. Basic Books.

Lewis, K., Kaufman, J., & Christakis, N. (2008). The taste for privacy: An analysis of college student privacy settings in an online social network. *Journal of Computer-Mediated Communication, 14*, 79–100.

Liu, Y., Kliman-Silver, C., & Mislove, A. (2014). The Tweets they are a-changin': Evolution of Twitter users and behavior.

Lukes, S. (1974). *Power: A radical view* (Vol. 1). London, England: London Macmillan.

Lynley, M. (2012, December 19). How to download your entire Twitter history (and laugh at how bad it is). *Wall Street Journal*. Retrieved from http://blogs.wsj.com/digits/2012/12/19/how-to-download-your-entire-twitter-history-and-laugh-at-how-bad-it-is/

Machlup, F. (1983). *Semantic quirks in studies of information. Machlup, F., & Mansfield, U. The Study of Information: Interdisciplinary Messages*. New York: Wiley.

Mackey, R. (2014, May 22). Twitter agrees to block "blasphemous" Tweets in Pakistan. *The New York Times*. Retrieved from http://www.nytimes.com/2014/05/22/world/asia/twitter-agrees-to-block-blasphemous-tweets-in-pakistan.html

Madden, M., Lenhart, A., Duggan, M., Cortesi, S., & Gasser, U. (2013). *Teens and technology 2013*. Pew Internet & American Life Project. Retrieved from http://www.pewinternet.org/2013/03/13/teens-and-technology-2013/

Mahmood, S., & Desmedt, Y. (2012). Usable privacy by visual and interactive control of information flow. In B. Christianson, J. Malcolm, F. Stajano, & J. Anderson (Eds.), *Security Protocols XX* (pp. 181–188). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-35694-0_20

Mansfield-Devine, S. (2011). Anonymous: Serious threat or mere annoyance? *Network Security*, *2011*(1), 4–10.

Marwick, A., & boyd, danah. (2011). To see and be seen: Celebrity practice on Twitter. *Convergence: The International Journal of Research into New Media Technologies*, *17*(2), 139 –158. http://doi.org/10.1177/1354856510394539

Marwick, A. E., & others. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, *13*(1), 114–133.

McLuhan, M. (1964). *Understanding media: The extensions of man*. Cambridge, MA: MIT Press.

Meece, M. (2012). President Obama's Consumer Privacy Bill of Rights. *Forbes.* Retrieved March 8, 2015, from http://www.forbes.com/sites/mickeymeece/2012/02/23/president-obamas-consumer-privacy-bill-of-rights/

Miller, V. (2008). New media, networking, and phatic culture. *Convergence*, *14(4)*, 387 – 400.

Milne, G. R., & Culnan, M. J. (2004). Strategies for reducing online privacy risks: Why consumers read (or don't read) online privacy notices. *Journal of Interactive Marketing*, *18*(3), 15–29.

Mitchell, A., & Guskin, E. (2013, November 4). Twitter news consumers: Young, mobile and educated. Retrieved from http://www.journalism.org/2013/11/04/twitter-news-consumers-young-mobile-and-educated/

Mokyr, J. (2004). *The gifts of Athena: Historical origins of the knowledge economy*. Princeton, NJ: Princeton University Press.

Morozov, E. (2012). *The net delusion: The dark side of internet freedom*. New York, NY: PublicAffairs.

Morozov, E. (2013). The meme hustler. *The Baffler*, (22). Retrieved from http://www.thebaffler.com/past/the_meme_hustler

Murphy, D. (2014). 44 percent of Twitter accounts have never tweeted. *PCMag.* Retrieved November 4, 2014, from http://www.pcmag.com/article2/0%2c2817%2c2456489%2c00.asp

National Telecommunications & Information Administration. (2012). White House unveils new comprehensive privacy blueprint. Retrieved from http://www.ntia.doc.gov/blog/2012/white-house-unveils-new-comprehensive-privacy-blueprint

Norman, D. A. (1983). Some observations on mental models. In D. Gentner & A. Stevens (Eds.), *Mental Models* (pp. 7–14). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Norman, D. A. (1988). *The psychology of everyday things*. New York, NY: Basic books.

Nye, J. S. (2008). Public diplomacy and soft power. *The Annals of the American Academy of Political and Social Science*, *616*(1), 94–109.

O'Reilly, T. (2005a). Web 2.0: Compact definition? *O'Reilly Media*. Retrieved from http://radar.oreilly.com/archives/2005/10/web-20-compact-definition.html

O'Reilly, T. (2005b). What is Web 2.0. *O'Reilly Media*. Retrieved August 30, 2013, from http://oreilly.com/web2/archive/what-is-web-20.html

Orlikowski, W. J., & Gash, D. C. (1994). Technological frames: Making sense of information technology in organizations. *ACM Transactions on Information Systems (TOIS)*, *12*(2), 174–207.

Park, Y. J. (2013). Digital literacy and privacy behavior online. *Communication Research*, *40*(2), 215–236. http://doi.org/10.1177/0093650211418338

Parsons, T. (1963). On the concept of political power. *Proceedings of the American Philosophical Society*, *107*(3), 232–262.

PCMag. (2014). Terms of service. *PCMag.* Retrieved January 6, 2015, from http://www.pcmag.com/encyclopedia/term/62682/terms-of-service

Petersen, S. (2008). Loser generated content: From participation to exploitation. *First Monday*, *13*(3). Retrieved from http://frodo.lib.uic.edu/ojsjournals/index.php/fm/article/view/2141

Pew Research Internet Project. (2014, January). Social networking fact sheet. Retrieved from http://www.pewinternet.org/fact-sheets/social-networking-fact-sheet/

Pfaffenberger, B. (1992). Technological dramas. *Science, Technology & Human Values*, *17*(3), 282–312.

Pinch, T. J., & Bijker, W. E. (1984). The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. *Social Studies of Science*, *14*(3), 399–441.

Popper, B. (2014). CNN announces partnership with Twitter to “revolutionize” news gathering. Retrieved April 27, 2014, from http://www.theverge.com/2014/1/29/5357826/cnn-announces-partnership-with-twitter

Potter, J. (1996). *Representing reality: Discourse, rhetoric and social construction*. Thousand Oaks, CA: SAGE.

Puschmann, C., & Burgess, J. (2013). The politics of Twitter data. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2206225

Quan-Haase, A., & Young, A. L. (2010). Uses and gratifications of social media: A comparison of Facebook and instant messaging. *Bulletin of Science, Technology & Society, 30*(5), 350–361.

Rapley, T. (2008). *Doing conversation, discourse and document analysis*. Thousand Oaks, CA: Sage Publications.

Raymond, M. (2010). How Tweet it is!: Library acquires entire Twitter archive. *Library of Congress Blog*. Retrieved November 11, 2011, from http://blogs.loc.gov/loc/2010/04/how-tweet-it-is-library-acquires-entire-twitter-archive/

Recuero, R., Amaral, A., & Monteiro, C. (2012). Fandoms, trending topics and social capital in Twitter. *Selected Papers of Internet Research*, *2*.

Reidenberg, J. R., Breaux, T., Cranor, L. F., French, B., Grannis, A., Graves, J. T., … Schaub, F. (2014). *Disagreeable privacy policies: Mismatches between meaning and users’ understanding* (SSRN Scholarly Paper No. ID 2418297). Rochester,

NY: Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=2418297

Ritzer, G., & Jurgenson, N. (2010). Production, consumption, prosumption. *Journal of Consumer Culture*, *10*(1), 13 –36. http://doi.org/10.1177/1469540509354673

Rogers, E. M. (2003). *Diffusion of Innovations* (5th ed.). New York, NY: Free Press.

Ronson, J. (2015, February 12). How one stupid Tweet blew up Justine Sacco's life. *The New York Times*. Retrieved from http://www.nytimes.com/2015/02/15/magazine/how-one-stupid-tweet-ruined-justine-saccos-life.html

Rooney, D. (1997). A contextualising, sociotechnical definition of technology: Learning from ancient Greece and Foucault. *Prometheus*, *15*(3), 399–407. http://doi.org/10.1080/08109029708632084

Roose, K. (2015). Meet the tweet-deleters: People who are making their Twitter histories self-destruct. *Fusion.* Retrieved from http://fusion.net/story/50322/meet-the-tweet-deleters-people-who-are-making-their-twitter-histories-self-destruct/

RT News. (2011). Censored: #Occupywallstreet. Retrieved November 8, 2014, from http://rt.com/news/twitter-ows-protest-censorship-653/

Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). Earthquake shakes Twitter users: Real-time event detection by social sensors. In *Proceedings of the 19th international conference on World wide web* (pp. 851–860).

Sanford, M. (2010). Growing around the world. *Twitter Blog.* Retrieved April 19, 2014, from https://blog.twitter.com/2010/growing-around-world

Scholz, T. (2008). Market ideology and the myths of Web 2.0. *First Monday*, *13(3)*. Retrieved from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2138/1945

Scott, J. C. (1990). *Domination and the arts of resistance: Hidden transcripts*. New Haven, CT: Yale University Press.

Semuels, A. (2009, February 20). Micro-blogging service Twitter dumps Dalai Lama impostor. *News India - Times*, p. 18. New York, N.Y., United States.

Seward, Z. (2013). What Twitter's IPO means. *Quartz.* Retrieved from http://qz.com/140197/what-twitters-ipo-means/

Shannon, C. E., & Weaver, W. (1949). *A mathematical model of communications*. Urbana, IL: University of Illinois Press.

Shapiro, J. J., & Hughes, S. K. (1996). Information literacy as a liberal art? *Educom Review*, *31*, 31–35.

Sherman, E. (2008). Privacy policies are great -- for PhDs. *CBS News*. Retrieved January 7, 2015, from http://www.cbsnews.com/news/privacy-policies-are-great-for-phds/

Shirky, C. (2011a). *Cognitive surplus: How technology makes consumers into collaborators* (Reprint edition). New York, NY: Penguin Books.

Shirky, C. (2011b). The political power of social media-technology, the public sphere, and political change. *Foreign Affairs*, *90*, 28.

Shrivastava, A. (2013). More relevant ads with tailored audiences. *Twitter Blogs*. Retrieved April 27, 2014, from https://blog.twitter.com/2013/more-relevant-ads-with-tailored-audiences

Shrum, L. J. (2002). Media consumption and perceptions of social reality: Effects and underlying processes. In J. Bryant & Oliver (Eds.), *Media effects: Advances in theory and research* (Vol. 2, pp. 69–95).

Shuen, A. (2008). *Web 2.0: A Strategy Guide: Business Thinking and Strategies Behind Successful Web 2.0 Implementations.* Sebastopol, CA: O'Reilly Media, Inc.

Signorini, A., Segre, A. M., & Polgreen, P. M. (2011). The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 Pandemic. *Plos One*, *6*(5). http://doi.org/10.1371/journal.pone.0019467

Siles, I. (2013). Inventing Twitter: An iterative approach to new media development. *International Journal of Communication*, *7*, 23.

Singel, R. (2010, April 14). Library of Congress archives Twitter history, while Google searches it. *Wired*. Retrieved November 11, 2011, from http://www.wired.com/epicenter/2010/04/loc-google-twitter/

Singletary, T. (2012). How do I get firehose access? *Twitter Developers Forum.* Retrieved April 4, 2014, from https://dev.twitter.com/discussions/2752

Slutsky, I., & Codel, E. (2006). *A look at the early days of Twitter* (Vol. 11). PodTech.

Small, H., Kasianovitz, K., Blanford, R., & Celaya, I. (2012). What your Tweets tell us about you: Identity, ownership and privacy of Twitter data. *International Journal of Digital Curation*, *7*(1), 174–197.

Sprenger, T. O., & Welpe, I. M. (2011). Tweets and peers: Defining industry groups and strategic peers based on investor perceptions of stocks on Twitter. *Algorithmic Finance*, *1*(1).

Statistic Brain. (2014). Twitter statistics. *Statistic Brain*. Retrieved April 12, 2014, from http://www.statisticbrain.com/twitter-statistics/

Stein, S. R. (2002). The "1984" Macintosh ad: Cinematic icons and constitutive rhetoric in the launch of a new machine. *Quarterly Journal of Speech*, *88*(2), 169–192.

Stone, B. (2008). An ongoing battle. *Twitter Blog.* Retrieved November 11, 2014, from https://blog.twitter.com/2008/ongoing-battle

Stone, B. (2009a). New front page! *Twitter Blog.* Retrieved January 6, 2015, from https://blog.twitter.com/2009/new-front-page

Stone, B. (2009b). Project Retweet: Phase one. *Twitter Blog.* Retrieved March 14, 2014, from https://blog.twitter.com/2009/project-retweet-phase-one

Stone, B. (2009c). Replies are now mentions. *Twitter Blog.* Retrieved April 12, 2014, from https://blog.twitter.com/2009/replies-are-now-mentions

Stone, B. (2009d). Retweet limited rollout. *Twitter Blog.* Retrieved April 12, 2014, from https://blog.twitter.com/2009/retweet-limited-rollout

Stone, B. (2010a). Hello world. *Twitter Blog.* Retrieved April 26, 2014, from https://blog.twitter.com/2010/hello-world

Stone, B. (2010b). Tweet preservation. *Twitter Blog.* Retrieved December 20, 2011, from http://blog.twitter.com/2010/04/tweet-preservation.html

Suh, B., Hong, L., Pirolli, P., & Chi, E. H. (2010). Want to be retweeted? Large scale analytics on factors impacting retweet in Twitter network. In *Social Computing (socialcom), 2010 IEEE Second International Conference* (pp. 177–184). IEEE.

Tapscott, D., & Williams, A. D. (2008). *Wikinomics: How mass collaboration changes everything*. New York, NY: Penguin.

TermsFeed. (2014). What is a Terms of Service agreement ? *Terms Feed*. Retrieved January 6, 2015, from https://termsfeed.com/resources/terms-service-agreement/

Terranova, T. (2000). Free labor: Producing culture for the digital economy. *Social Text*, (63), 33–58.

Terranova, T. (2004). *Network culture: Politics for the information age*. London, England: Pluto Press.

Thorson, K., Driscoll, K., Ekdale, B., Edgerly, S., Thompson, L. G., Schrock, A., ... & Wells, C. (2013). YouTube, Twitter and the Occupy movement: Connecting content and circulation practices. *Information, Communication & Society*, *16*(3), 421-451.

Tiernan, P. (2013). A study of the use of Twitter by students for lecture engagement and discussion. *Education and Information Technologies*, 1–18.

Toffler, A. (1984). *The Third Wave*. New York, NY: Bantam.

Tsukayama, H. (2012, January 27). Twitter's country-specific censorship tool prompts user protest. *The Washington Post*. Retrieved from http://www.washingtonpost.com/business/technology/twitters-country-specific-censorship-tool-prompts-user-protest/2012/01/27/gIQALWoMVQ_story.html

Tsukayama, H. (2013). Twitter turns 7: Users send over 400 million tweets per day. *The Washington Post.* Retrieved August 30, 2013, from http://articles.washingtonpost.com/2013-03-21/business/37889387_1_tweets-jack-dorsey-twitter

Tufekci, Z. (2008). Can you see me now? Audience and disclosure regulation in online social network sites. *Bulletin of Science, Technology & Society*, *28*(1), 20–36.

Tuominen, K. (1997). User-centered discourse: An analysis of the subject positions of the user and the librarian. *The Library Quarterly*, *67*(4), 350–371.

Turow, J. (2003). *Americans & online privacy: The system is broken*. Annenberg Public Policy Center, University of Pennsylvania.

Turow, J., Feldman, L., & Meltzer, K. (2005). Open to Exploitation: America's Shoppers Online and Offline: A Report from the Annenberg Public Policy Center of the University of Pennsylvania. *Annenberg School for Communication Departmental Papers (ASC)*.

twitter.com. (2009, January 18). The Twitter Rules. *Twitter Support.* Retrieved January 7, 2015, from https://web.archive.org/web/20090118211301/http://twitter.zendesk.com/forums/26257/entries/18311

twitter.com. (2011). Let's fly. *Twitter Blog.* Retrieved January 22, 2015, from https://blog.twitter.com/2011/lets-fly

twitter.com. (2012). History of the REST & Search API. *Twitter Developers*. Retrieved April 5, 2014, from https://dev.twitter.com/docs/history-rest-search-api

twitter.com. (2012). Twitter Terms of Service. *Twitter.com.* Retrieved June 14, 2012, from https://twitter.com/tos

twitter.com. (2013a). Developer rules of the road. *Twitter Developers*. Retrieved April 5, 2014, from https://dev.twitter.com/terms/api-terms

twitter.com. (2013b). New compete study: Primary mobile users on Twitter. *Twitter Blog.* Retrieved November 4, 2014, from https://blog.twitter.com/2013/new-compete-study-primary-mobile-users-on-twitter

twitter.com. (2013c). Overview: Version 1.1 of the Twitter API. *Twitter Developers*. Retrieved April 4, 2014, from https://dev.twitter.com/docs/api/1.1/overview

twitter.com. (2013d). Tweet Button. *Twitter Developers*. Retrieved April 7, 2014, from https://dev.twitter.com/docs/tweet-button

twitter.com. (2013e). Twitter Privacy Policy. *Twitter.com.* Retrieved February 21, 2014, from https://twitter.com/privacy

twitter.com. (2013f). Using the Twitter Search API. *Twitter Developers*. Retrieved April 4, 2014, from https://dev.twitter.com/docs/using-search

twitter.com. (2013, October 26). Twitter's use of cookies and similar technologies. *Twitter Help Center*. Retrieved January 21, 2015, from https://web.archive.org/web/20131026202301/https://support.twitter.com/articles/20170514-twitters-use-of-cookies-and-similar-technologies

twitter.com. (2014a). A field guide to Twitter Platform objects. *Twitter Developers.* Retrieved January 21, 2014, from https://dev.twitter.com/docs/platform-objects

twitter.com. (2014b). Analytics. *Twitter Business Solutions*. Retrieved August 19, 2014, from https://business.twitter.com/products/analytics

twitter.com. (2014c). DataSift. *Twitter Certified Products.* Retrieved January 23, 2015, from https://biz.twitter.com/partners/certified-products/datasift

twitter.com. (2014d). Embedded Timelines. *Twitter Developers*. Retrieved April 4, 2014, from https://dev.twitter.com/docs/embedded-timelines

twitter.com. (2014e). Entities. *Twitter Developers*. Retrieved January 21, 2014, from https://dev.twitter.com/docs/platform-objects/entities

twitter.com. (2014f). Find a partner: Certified Products. *Twitter for Business*. Retrieved April 20, 2014, from https://business.twitter.com/partners/list/certified-products

twitter.com. (2014g). Frequently Asked Questions. *Twitter Developers*. Retrieved January 21, 2014, from https://dev.twitter.com/docs/faq

twitter.com. (2014h). Places. *Twitter Developers*. Retrieved January 21, 2014, from https://dev.twitter.com/docs/platform-objects/places

twitter.com. (2014i). REST API Rate Limiting in v1.1. *Twitter Developers*. Retrieved April 4, 2014, from https://dev.twitter.com/docs/rate-limiting/1.1

twitter.com. (2014j). SocialBro. *Twitter for Businesses.* Retrieved January 23, 2015, from https://biz.twitter.com/partners/socialbro

twitter.com. (2014k). The Streaming APIs. *Twitter Developers*. Retrieved January 21, 2014, from https://dev.twitter.com/docs/streaming-apis

twitter.com. (2014l). Things Every Developer Should Know. *Twitter Developers.* Retrieved April 5, 2014, from https://dev.twitter.com/docs/things-every-developer-should-know

twitter.com. (2014m). Tweets. *Twitter Developers*. Retrieved January 21, 2014, from https://dev.twitter.com/docs/platform-objects/tweets

twitter.com. (2014n). Twitter Buttons. *About Twitter*. Retrieved May 12, 2014, from https://about.twitter.com/resources/buttons

twitter.com. (2014o). Twitter Certified Products Program. *Twitter Developers*. Retrieved January 21, 2014, from https://dev.twitter.com/programs/twitter-certified-products

twitter.com. (2014p). About tailored suggestions. *Twitter Help Center.* Retrieved April 11, 2014, from https://support.twitter.com/articles/20169421-about-tailored-suggestions#

twitter.com. (2014q). FAQs about following. *Twitter Help Center.* Retrieved April 12, 2014, from https://support.twitter.com/groups/52-notifications/topics/213-following/articles/14019-faqs-about-following#

twitter.com. (2014r). FAQs about Retweets (RT). *Twitter Help Center.* Retrieved April 12, 2014, from https://support.twitter.com/articles/77606-faqs-about-retweets-rt#

twitter.com. (2014s). FAQs about Trends on Twitter. *Twitter Help Center.* Retrieved April 3, 2014, from http://support.twitter.com/articles/101125-faqs-about-trends-on-twitter#

twitter.com. (2014t). FAQs about verified accounts. *Twitter Help Center.* Retrieved April 19, 2014, from https://support.twitter.com/articles/119135-faqs-about-verified-accounts#

twitter.com. (2014u). Favoriting a Tweet. *Twitter Help Center.* Retrieved April 12, 2014, from https://support.twitter.com/articles/20169874-favoriting-a-tweet#

twitter.com. (2014v). Media settings and best practices. *Twitter Help Center.* Retrieved April 11, 2014, from http://support.twitter.com/articles/20169200-media-settings-and-best-practices#

twitter.com. (2014w). Posting or deleting direct messages. *Twitter Help Center.* Retrieved April 12, 2014, from https://support.twitter.com/articles/14606#

twitter.com. (2014x). The Twitter Rules. *Twitter Help Center.* Retrieved April 2, 2014, from https://support.twitter.com/articles/18311-the-twitter-rules#

twitter.com. (2014y). Twitter's suggestions for who to follow. *Twitter Help Center.* Retrieved April 17, 2014, from https://support.twitter.com/articles/227220-twitter-s-suggestions-for-who-to-follow#

twitter.com. (2014z). Twitter Ads policies. *Twitter Help Center.* Retrieved April 29, 2014, from https://support.twitter.com/adspolicy#

twitter.com. (2014aa). What are Promoted Tweets? *Twitter Help Center.* Retrieved April 19, 2014, from https://support.twitter.com/articles/142101-what-are-promoted-tweets#

twitter.com. (2014ab). What are @replies and mentions? *Twitter Help Center.* Retrieved April 12, 2014, from https://support.twitter.com/groups/52-notifications/topics/212-replies-mentions/articles/14023-what-are-replies-and-mentions#

twitter.com. (2014ac). Your privacy controls for tailored ads. *Twitter Help Center.* Retrieved April 11, 2014, from https://support.twitter.com/articles/20170405#

twitter.com. (2014ad). Twitter Terms of Service. *Twitter.* Retrieved February 21, 2014, from https://twitter.com/tos

twitter.com. (2014ae). Users. *Twitter Developers.* Retrieved January 21, 2014, from https://dev.twitter.com/docs/platform-objects/users

Twitter, Inc. (2013). *Form S-1* (Form S-1). Washington, D.C.: Securities and Exchange Commission. Retrieved from http://edgar.sec.gov/Archives/edgar/data/1418091/000119312513390321/d564001ds1.htm

U. S. Census Bureau, D. I. S. (2013). Educational attainment in the United States: 2013 - detailed tables. Retrieved November 4, 2014, from https://www.census.gov/hhes/socdemo/education/data/cps/2013/tables.html

Van Dijck, J. (2009). Users like you? Theorizing agency in user-generated content. *Media, Culture, and Society*, *31*(1), 41.

Van Dijck, J. (2011). Tracing Twitter: The rise of a microblogging platform. *International Journal of Media & Cultural Politics*, *7*(3), 333–348.

Van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. Oxford, England: Oxford Univ. Press.

Van Dijck, J., & Nieborg, D. (2009). Wikinomics and its discontents: a critical analysis of Web 2.0 business manifestos. *New Media & Society*, *11*(5), 855–874.

Van Dijk, T. A. (1989). Structures of discourse and structures of power. *Communication Yearbook*, *12*, 18–59.

Van Dijk, T. A. (1993). Principles of critical discourse analysis. *Discourse & Society*, *4*(2), 249–283.

Van Dijk, T. A. (1996). Discourse, power and access. *Texts and Practices: Readings in Critical Discourse Analysis*, 84–104.

Van Dijk, T. A. (2003). Critical discourse analysis. *The Handbook of Discourse Analysis*, *18*, 352.

Vitak, J. (2012). *Keeping connected in the Facebook age: The relationship between Facebook use, relationship maintenance strategies, and relational outcomes*. (Doctoral Dissertation). Michigan State University.

Wang, Y., Norcie, G., Komanduri, S., Acquisti, A., Leon, P. G., & Cranor, L. F. (2011). I regretted the minute I pressed share: A qualitative study of regrets on Facebook. In *Proceedings of the Seventh Symposium on Usable Privacy and Security* (p. 10). ACM.

Waters, R. D., & Jamal, J. Y. (2011). Tweet, tweet, tweet: A content analysis of nonprofit organizations' Twitter updates. *Public Relations Review*, *37*(3), 321–324. http://doi.org/10.1016/j.pubrev.2011.03.002

Watson, B. R. (2012). *Is Twitter a counter public?: Comparing individual and community forces that shaped local Twitter and newspaper coverage of the BP oil spill (Doctoral Dissertation)*. University of North Carolina.

Westman, S., & Freund, L. (2010). Information interaction in 140 characters or less: Genres on Twitter. In *Proceedings of the Third Symposium on Information Interaction in Context* (pp. 323–328). New York, NY, USA: ACM. http://doi.org/10.1145/1840784.1840833

Wigand, F. D. L. (2010). Twitter in government: Building relationships one tweet at a time. In *Information Technology: New Generations (ITNG), 2010 Seventh International Conference on* (pp. 563–567). IEEE.

Williams, E. (2008). How @replies work on Twitter (and how they might). *Twitter Blogs.* Retrieved April 12, 2014, from https://blog.twitter.com/2008/how-replies-work-twitter-and-how-they-might

Williams, E. (2010a). #newtwitterceo. *Twitter Blogs*. Retrieved April 27, 2014, from https://blog.twitter.com/2010/newtwitterceo

Williams, E. (2010b). The Evolving Ecosystem. *Twitter Blog.* Retrieved January 8, 2015, from https://blog.twitter.com/2010/evolving-ecosystem

Windahl, S., Signitzer, B., & Olson, J. T. (2008). *Using Communication Theory: An Introduction to Planned Communication*. New York, NY: SAGE.

Wyatt, S. (2005). Non-users also matter: The construction of users and non-users of the Internet. In N. Oudshoorn, T. Pinch (Eds.), *How users matter. The co-construction of users and technology*, Cambridge, MA: MIT Press, pp. 67–79

YouTube.com. (2013). Statistics. *YouTube*. Retrieved August 30, 2013, from https://www.youtube.com/yt/press/statistics.html

Zhang, M., Jansen, B. J., & Chowdhury, A. (2011). Business engagement on Twitter: a path analysis. *Electronic Markets*, *21*(3), 161–175. http://doi.org/10.1007/s12525-011-0065-z

Zhang, X., Fuehres, H., & Gloor, P. A. (2011). Predicting Stock Market Indicators Through Twitter "I hope it is not as bad as I fear." *Procedia - Social and Behavioral Sciences*, *26*, 55–62. http://doi.org/10.1016/j.sbspro.2011.10.562

Zimmer, M., & Proferes, N. J. (2014). A topology of Twitter research: Disciplines, methods, and ethics. *Aslib Journal of Information Management*, *66*, 250–261.

Zubiaga, A., Spina, D., Fresno, V., & Martínez, R. (2011). Classifying trending topics: A typology of conversation triggers on twitter. In *Proceedings of the 20th ACM international conference on Information and knowledge management* (pp. 2461–2464). Retrieved from http://dl.acm.org/citation.cfm?id=2063992

Appendix A: REST API Methods and General Data Types

REST API Methods

Data Type	Resource	Description
Direct Message	GET direct_messages	Returns the 20 most recent direct messages sent to the authenticating user. Includes detailed information about the sender and recipient user. You can request up to 200 direct messages per call, up to a maximum of 800 incoming DMs. Important: This method requires an access token with RWD
Direct Message	GET direct_messages/sent	Returns the 20 most recent direct messages sent by the authenticating user. Includes detailed information about the sender and recipient user. You can request up to 200 direct messages per call, up to a maximum of 800 outgoing DMs. Important: This method requires an access token with RWD
Direct Message	GET direct_messages/show	Returns a single direct message, specified by an id parameter. Like the /1.1/direct_messages.format request, this method will include the user objects of the sender and recipient. Important: This method requires an access token with RWD
Direct Message	POST direct_messages/destroy	Destroys the direct message specified in the required ID parameter. The authenticating user must be the recipient of the specified direct message. Important: This method requires an access token with RWD
Direct Message	POST direct_messages/new	Sends a new direct message to the specified user from the authenticating user. Requires both the user and text parameters and must be a POST. Returns the sent message in the requested format if successful.
Favorites	GET favorites/list	Returns the 20 most recent Tweets favorited by the authenticating or specified user.
Favorites	POST favorites/destroy	Un-favorites the status specified in the ID parameter as the authenticating user. Returns the un-favorited status in the requested format when successful. This process invoked by this method is asynchronous. The immediately returned status may not indicate the resultant favorited status of the...
Favorites	POST favorites/create	Favorites the status specified in the ID parameter as the authenticating user. Returns the favorite status when successful. This process invoked by this method is asynchronous. The immediately returned status may not indicate the resultant favorited status of the tweet. A 200 OK response from this...

Data Type	Resource	Description
Friends and Followers	GET friendships/no_retweets/ids	Returns a collection of user_ids that the currently authenticated user does not want to receive retweets from. Use POST friendships/update to set the "no retweets" status for a given user account on behalf of the current user.
Friends and Followers	GET friends/ids	Returns a cursored collection of user IDs for every user the specified user is following (otherwise known as their "friends"). At this time, results are ordered with the most recent following first — however, this ordering is subject to unannounced change and eventual consistency issues....
Friends and Followers	GET followers/ids	Returns a cursored collection of user IDs for every user following the specified user. At this time, results are ordered with the most recent following first — however, this ordering is subject to unannounced change and eventual consistency issues. Results are given in groups of 5,000 user...
Friends and Followers	GET friendships/incoming	Returns a collection of numeric IDs for every user who has a pending request to follow the authenticating user.
Friends and Followers	GET friendships/outgoing	Returns a collection of numeric IDs for every protected user for whom the authenticating user has a pending follow request.
Friends and Followers	POST friendships/create	Allows the authenticating users to follow the user specified in the ID parameter. Returns the befriended user in the requested format when successful. Returns a string describing the failure condition when unsuccessful. If you are already friends with the user a HTTP 403 may be returned, though for...
Friends and Followers	POST friendships/destroy	Allows the authenticating user to unfollow the user specified in the ID parameter. Returns the unfollowed user in the requested format when successful. Returns a string describing the failure condition when unsuccessful. Actions taken in this method are asynchronous and changes will be eventually...
Friends and Followers	POST friendships/update	Allows one to enable or disable retweets and device notifications from the specified user.
Friends and Followers	GET friendships/show	Returns detailed information about the relationship between two arbitrary users.
Friends and Followers	GET friends/list	Returns a cursored collection of user objects for every user the specified user is following (otherwise known as their "friends"). At this time, results are ordered with the most recent following first — however, this ordering is subject to unannounced change and eventual consistency issues...

Data Type	Resource	Description
Friends and Followers	GET followers/list	Returns a cursored collection of user objects for users following the specified user. At this time, results are ordered with the most recent following first — however, this ordering is subject to unannounced change and eventual consistency issues. Results are given in groups of 20 users and...
Friends and Followers	GET friendships/lookup	Returns the relationships of the authenticating user to the comma-separated list of up to 100 screen_names or user_ids provided. Values for connections can be: following, following_requested, followed_by, none, blocking.
Help	GET help/configuration	Returns the current configuration used by Twitter including twitter.com slugs which are not usernames, maximum photo resolutions, and t.co URL lengths. It is recommended applications request this endpoint when they are loaded, but no more than once a day.
Help	GET help/languages	Returns the list of languages supported by Twitter along with their ISO 639-1 code. The ISO 639-1 code is the two letter value to use if you include lang with any of your requests.
Help	GET help/privacy	Returns Twitter's Privacy Policy.
Help	GET help/tos	Returns the Twitter Terms of Service in the requested format. These are not the same as the Developer Rules of the Road.
Help	GET application/rate_limit_status	Returns the current rate limits for methods belonging to the specified resource families. Each 1.1 API resource belongs to a "resource family" which is indicated in its method documentation. You can typically determine a method's resource family from the first component of the path after the...
Lists	GET lists/list	Returns all lists the authenticating or specified user subscribes to, including their own. The user is specified using the user_id or screen_name parameters. If no user is given, the authenticating user is used. This method used to be GET lists in version 1.0 of the API and has been renamed for...
Lists	GET lists/statuses	Returns a timeline of tweets authored by members of the specified list. Retweets are included by default. Use the include_rts=false parameter to omit retweets. Embedded Timelines is a great way to embed list timelines on your website.
Lists	POST lists/members/destroy	Removes the specified member from the list. The authenticated user must be the list's owner to remove members from the list.
Lists	GET lists/memberships	Returns the lists the specified user has been added to. If user_id or screen_name are not provided the memberships for the authenticating user are returned.

Data Type	Resource	Description
Lists	GET lists/subscribers	Returns the subscribers of the specified list. Private list subscribers will only be shown if the authenticated user owns the specified list.
Lists	POST lists/subscribers/create	Subscribes the authenticated user to the specified list.
Lists	GET lists/subscribers/show	Check if the specified user is a subscriber of the specified list. Returns the user if they are subscriber.
Lists	POST lists/subscribers/destroy	Unsubscribes the authenticated user from the specified list.
Lists	POST lists/members/create_all	Adds multiple members to a list, by specifying a comma-separated list of member ids or screen names. The authenticated user must own the list to be able to add members to it. Note that lists can't have more than 5,000 members, and you are limited to adding up to 100 members to a list at a time with...
Lists	GET lists/members/show	Check if the specified user is a member of the specified list.
Lists	GET lists/members	Returns the members of the specified list. Private list members will only be shown if the authenticated user owns the specified list.
Lists	POST lists/members/create	Add a member to a list. The authenticated user must own the list to be able to add members to it. Note that lists cannot have more than 5,000 members.
Lists	POST lists/destroy	Deletes the specified list. The authenticated user must own the list to be able to destroy it.
Lists	POST lists/update	Updates the specified list. The authenticated user must own the list to be able to update it.
Lists	POST lists/create	Creates a new list for the authenticated user. Note that you can't create more than 20 lists per account.
Lists	GET lists/show	Returns the specified list. Private lists will only be shown if the authenticated user owns the specified list.
Lists	GET lists/subscriptions	Obtain a collection of the lists the specified user is subscribed to, 20 lists per page by default. Does not include the user's own lists.
Lists	POST lists/members/destroy_all	Removes multiple members from a list, by specifying a comma-separated list of member ids or screen names. The authenticated user must own the list to be able to remove members from it. Note that lists can't have more than 500 members, and you are limited to removing up to 100 members to a list at a...
Lists	GET lists/ownerships	Returns the lists owned by the specified Twitter user. Private lists will only be shown if the authenticated user is also the owner of the lists.
Location Data	GET geo/id/:place_id	Returns all the information about a known place.

Data Type	Resource	Description
Location Data	GET geo/reverse_geocode	Given a latitude and a longitude, searches for up to 20 places that can be used as a place_id when updating a status. This request is an informative call and will deliver generalized results about geography.
Location Data	GET geo/search	Search for places that can be attached to a statuses/update. Given a latitude and a longitude pair, an IP address, or a name, this request will return a list of all the valid places that can be used as the place_id when updating a status. Conceptually, a query can be made from the user's location...
Location Data	GET geo/similar_places	Locates places near the given coordinates which are similar in name.
Location Data	POST geo/place	As of December 2nd, 2013, this endpoint is deprecated and retired and no longer functions. Place creation was used infrequently by third party applications and is generally no longer supported on Twitter. Requests will return with status 410 (Gone) with error code 251. Follow the discussion about...
Oauth	GET oauth/authenticate	Allows a Consumer application to use an OAuth request_token to request user authorization. This method is a replacement of Section 6.2 of the OAuth 1.0 authentication flow for applications using the callback authentication flow. The method will use the currently logged in user as the account for...
Oauth	GET oauth/authorize	Allows a Consumer application to use an OAuth Request Token to request user authorization. This method fulfills Section 6.2 of the OAuth 1.0 authentication flow. Desktop applications must use this method (and cannot use GET oauth/authenticate). Please use HTTPS for this method, and all other OAuth...
Oauth	POST oauth/access_token	Allows a Consumer application to exchange the OAuth Request Token for an OAuth Access Token. This method fulfills Section 6.3 of the OAuth 1.0 authentication flow. The OAuth access token may also be used for xAuth operations. Please use HTTPS for this method, and all other OAuth token negotiation...
Oauth	POST oauth/request_token	Allows a Consumer application to obtain an OAuth Request Token to request user authorization. This method fulfills Section 6.1 of the OAuth 1.0 authentication flow. It is strongly recommended you use HTTPS for all OAuth authorization steps. Usage Note: Only ASCII values are accepted for the...

Data Type	Resource	Description
Oauth	POST oauth2/token	Allows a registered application to obtain an OAuth 2 Bearer Token, which can be used to make API requests on an application's own behalf, without a user context. This is called Application-only authentication. A Bearer Token may be invalidated using oauth2/invalidate_token. Once a Bearer Token has...
Oauth	POST oauth2/invalidate_token	Allows a registered application to revoke an issued OAuth 2 Bearer Token by presenting its client credentials. Once a Bearer Token has been invalidated, new creation attempts will yield a different Bearer Token and usage of the invalidated token will no longer be allowed. As with all API v1.1...
Report Spam	POST users/report_spam	Report the specified user as a spam account to Twitter. Additionally performs the equivalent of POST blocks/create on behalf of the authenticated user.
Saved Searches	GET saved_searches/list	Returns the authenticated user's saved search queries.
Saved Searches	GET saved_searches/show/:id	Retrieve the information for the saved search represented by the given id. The authenticating user must be the owner of saved search ID being requested.
Saved Searches	POST saved_searches/create	Create a new saved search for the authenticated user. A user may only have 25 saved searches.
Saved Searches	POST saved_searches/destroy/:id	Destroys a saved search for the authenticating user. The authenticating user must be the owner of saved search id being destroyed.
Search	Resource	Description
Search	GET search/tweets	Returns a collection of relevant Tweets matching a specified query. Please note that Twitter's search service and, by extension, the Search API is not meant to be an exhaustive source of Tweets. Not all Tweets will be indexed or made available via the search interface. In API v1.1, the response...
Search	Streaming	
Search	Resource	Description
Search	POST statuses/filter	Returns public statuses that match one or more filter predicates. Multiple parameters may be specified which allows most clients to use a single connection to the Streaming API. Both GET and POST requests are supported, but GET requests with too many parameters may cause the request to be...
Search	GET statuses/sample	Returns a small random sample of all public statuses. The Tweets returned by the default access level are the same, so if two different clients connect to this endpoint, they will see the same Tweets.

Data Type	Resource	Description
Search	GET statuses/firehose	This endpoint requires special permission to access. Returns all public statuses. Few applications require this level of access. Creative use of a combination of other resources and various access levels can satisfy nearly every application use case.
Search	GET user	Streams messages for a single user, as described in User streams.
Search	GET site	Streams messages for a set of users, as described in Site streams.
Suggested Users	GET users/suggestions/:slug	Access the users in a given category of the Twitter suggested user list. It is recommended that applications cache this data for no more than one hour.
Suggested Users	GET users/suggestions	Access to Twitter's suggested user list. This returns the list of suggested user categories. The category can be used in GET users/suggestions/:slug to get the users in that category.
Suggested Users	GET users/suggestions/:slug/members	Access the users in a given category of the Twitter suggested user list and return their most recent status if they are not a protected user.
Timelines	GET statuses/mentions_timeline	Returns the 20 most recent mentions (tweets containing a users's @screen_name) for the authenticating user. The timeline returned is the equivalent of the one seen when you view your mentions on twitter.com. This method can only return up to 800 tweets. See Working with Timelines for...
Timelines	GET statuses/user_timeline	Returns a collection of the most recent Tweets posted by the user indicated by the screen_name or user_id parameters. User timelines belonging to protected users may only be requested when the authenticated user either "owns" the timeline or is an approved follower of the owner. The timeline...
Timelines	GET statuses/home_timeline	Returns a collection of the most recent Tweets and retweets posted by the authenticating user and the users they follow. The home timeline is central to how most users interact with the Twitter service. Up to 800 Tweets are obtainable on the home timeline. It is more volatile for users that follow...
Timelines	GET statuses/retweets_of_me	Returns the most recent tweets authored by the authenticating user that have been retweeted by others. This timeline is a subset of the user's GET statuses/user_timeline. See Working with Timelines for instructions on traversing timelines.
Trends	GET trends/place	Returns the top 10 trending topics for a specific WOEID, if trending information is available for it. The response is an array of "trend" objects that encode the name of the trending topic, the query parameter that can be used to search for the topic on Twitter Search, and the Twitter Search URL....

Data Type	Resource	Description
Trends	GET trends/available	Returns the locations that Twitter has trending topic information for. The response is an array of "locations" that encode the location's WOEID and some other human-readable information such as a canonical name and country the location belongs in. A WOEID is a Yahoo! Where On Earth ID.
Trends	GET trends/closest	Returns the locations that Twitter has trending topic information for, closest to a specified location. The response is an array of "locations" that encode the location's WOEID and some other human-readable information such as a canonical name and country the location belongs in. A WOEID is a Yahoo...
Tweets	GET statuses/retweets/:id	Returns a collection of the 100 most recent retweets of the tweet specified by the id parameter.
Tweets	GET statuses/show/:id	Returns a single Tweet, specified by the id parameter. The Tweet's author will also be embedded within the tweet. See Embeddable Timelines, Embeddable Tweets, and GET statuses/oembed for tools to render Tweets according to Display Requirements.
Tweets	POST statuses/destroy/:id	Destroys the status specified by the required ID parameter. The authenticating user must be the author of the specified status. Returns the destroyed status if successful.
Tweets	POST statuses/update	Updates the authenticating user's current status, also known as tweeting. To upload an image to accompany the tweet, use POST statuses/update_with_media. For each update attempt, the update text is compared with the authenticating user's recent tweets. Any attempt that would result in duplication...
Tweets	POST statuses/retweet/:id	Retweets a tweet. Returns the original tweet with retweet details embedded.
Tweets	POST statuses/update_with_media	Updates the authenticating user's current status and attaches media for upload. In other words, it creates a Tweet with a picture attached. Unlike POST statuses/update, this method expects raw multipart data. Your POST request's Content-Type should be set to multipart/form-data with the media[]...
Tweets	GET statuses/oembed	Returns information allowing the creation of an embedded representation of a Tweet on third party sites. See the oEmbed specification for information about the response format. While this endpoint allows a bit of customization for the final appearance of the embedded Tweet, be aware that the...
Tweets	GET statuses/retweeters/ids	Returns a collection of up to 100 user IDs belonging to users who have retweeted the tweet specified by the id parameter. This method offers similar data to GET statuses/retweets/:id and replaces API v1's GET statuses/:id/retweeted_by/ids method.

Data Type	Resource	Description
Users	GET account/settings	Returns settings (including current trend, geo and sleep time information) for the authenticating user.
Users	GET account/verify_credentials	Returns an HTTP 200 OK response code and a representation of the requesting user if authentication was successful; returns a 401 status code and an error message if not. Use this method to test if supplied user credentials are valid.
Users	POST account/settings	Updates the authenticating user's settings.
Users	POST account/update_delivery_device	Sets which device Twitter delivers updates to for the authenticating user. Sending none as the device parameter will disable SMS updates.
Users	POST account/update_profile	Sets values that users are able to set under the "Account" tab of their settings page. Only the parameters specified will be updated.
Users	POST account/update_profile_background_image	Updates the authenticating user's profile background image. This method can also be used to enable or disable the profile background image. Although each parameter is marked as optional, at least one of image, tile or use must be provided when making this request.
Users	POST account/update_profile_colors	Sets one or more hex values that control the color scheme of the authenticating user's profile page on twitter.com. Each parameter's value must be a valid hexidecimal value, and may be either three or six characters (ex: #fff or #ffffff).
Users	POST account/update_profile_image	Updates the authenticating user's profile image. Note that this method expects raw multipart data, not a URL to an image. This method asynchronously processes the uploaded file before updating the user's profile image URL. You can either update your local cache the next time you request the user's...
Users	GET blocks/list	Returns a collection of user objects that the authenticating user is blocking. Important On October 15, 2012 this method will become cursored by default, altering the default response format. See Using cursors to navigate collections for more details on how cursoring works.
Users	GET blocks/ids	Returns an array of numeric user ids the authenticating user is blocking. Important On October 15, 2012 this method will become cursored by default, altering the default response format. See Using cursors to navigate collections for more details on how cursoring works.
Users	POST blocks/create	Blocks the specified user from following the authenticating user. In addition the blocked user will not show in the authenticating users mentions or timeline (unless retweeted by another user). If a follow or friend relationship exists it is destroyed.

Data Type	Resource	Description
Users	POST blocks/destroy	Un-blocks the user specified in the ID parameter for the authenticating user. Returns the un-blocked user in the requested format when successful. If relationships existed before the block was instated, they will not be restored.
Users	GET users/lookup	Returns fully-hydrated user objects for up to 100 users per request, as specified by comma-separated values passed to the user_id and/or screen_name parameters. This method is especially useful when used in conjunction with collections of user IDs returned from GET friends/ids and GET followers/...
Users	GET users/show	Returns a variety of information about the user specified by the required user_id or screen_name parameter. The author's most recent Tweet will be returned inline when possible. GET users/lookup is used to retrieve a bulk collection of user objects.
Users	GET users/search	Provides a simple, relevance-based search interface to public user accounts on Twitter. Try querying by topical interest, full name, company name, location, or other criteria. Exact match searches are not supported. Only the first 1,000 matching results are available.
Users	GET users/contributees	Returns a collection of users that the specified user can "contribute" to.
Users	GET users/contributors	Returns a collection of users who can contribute to the specified account.
Users	POST account/remove_profile_banner	Removes the uploaded profile banner for the authenticating user. Returns HTTP 200 upon success.
Users	POST account/update_profile_banner	Uploads a profile banner on behalf of the authenticating user. For best results, upload an
Users	GET users/profile_banner	Returns a map of the available size variations of the specified user's profile banner. If the user has not uploaded a profile banner, a HTTP 404 will be served instead. This method can be used instead of string manipulation on the profile_banner_url returned in user objects as described in User...

REST API General Data Types

Data Type	Description
Direct Messages	Direct Messages are short, non-public messages sent between two users. Access to Direct Messages is governed by the The Application Permission Model.
Favorites	Users favorite tweets to give recognition to awesome tweets, to curate the best of Twitter, to save for reading later, and a variety of other reasons. Likewise, developers make use of "favs" in many different ways.
Followers and Friends	Users follow their interests on Twitter through both one-way and mutual following relationships.
Help	These methods assist you in working & debugging with the Twitter API.
Lists	Lists are collections of tweets, culled from a curated list of Twitter users. List timeline methods include tweets by all members of a list.
OAuth	Twitter uses OAuth for authentication. Be sure and read about Authentication & Authorization.
Places & Geo	Users tweet from all over the world. These methods allow you to attach location data to tweets and discover tweets & locations.
Saved Searches	Allows users to save references to search criteria for reuse later.
Search	Find relevant Tweets based on queries performed by your users.
Spam Reporting	These methods are used to report user accounts as spam accounts.
Streaming	No description given.
Suggested Users	Categorical organization of users that others may be interested to follow.
Timelines	Timelines are collections of Tweets, ordered with the most recent first.
Trends	With so many tweets from so many users, themes are bound to arise from the zeitgeist. The Trends methods allow you to explore what's trending on Twitter.
Tweets	Tweets are the atomic building blocks of Twitter, 140-character status updates with additional associated metadata. People tweet for a variety of reasons about a multitude of topics.
Users	Users are at the center of everything Twitter: they follow, they favorite, and tweet & retweet.

Appendix B: Mapping of Specific Data Types Mentioned in the Privacy Policy

code	Privacy Policy Category	General Classification of Data (Author Generated)	Detailed Type of Data
a1	"Information Collected Upon Registration"	User Account Info	Name
a2	"Information Collected Upon Registration"	User Account Info	Username
a3	"Information Collected Upon Registration"	User Account Info	Password
a4	"Information Collected Upon Registration"	User Account Info	E-Mail Address
b1	"Additional Information"	User Profile Info	Short Biography
b2	"Additional Information"	User Profile Info	Your Location (profile)
b3	"Additional Information"	User Profile Info	Your Website
b4	"Additional Information"	User Profile Info	Picture
b5	"Additional Information"	User Customized Account Info	Cellphone Number for SMS Delivery
b6	"Additional Information"	User Customized Account Info	Address Book information
b7	"Additional Information"	User Customized Account Info	Linked Services
b8	"Additional Information"	User Customized Account Info	Registration or Profile Information from Linked Services
c1	"Tweets, Following, and other Public Information"	User Tweet	140 characters
c2	"Tweets, Following, and other Public Information"	User Tweet	Content, "link photos, videos, and links"
c3	"Tweets, Following, and other Public Information"	User Tweet Metadata	Category: Tweet Metadata
c4	"Tweets, Following, and other Public Information"	User Tweet Metadata	When you tweeted
c5	"Tweets, Following, and other Public Information"	User Metadata	Lists you create

code	Privacy Policy Category	General Classification of Data (Author Generated)	Detailed Type of Data
c6	"Tweets, Following, and other Public Information"	User Metadata	People you Follow
c7	"Tweets, Following, and other Public Information"	User Metadata	Tweets you mark as Favorite
c8	"Tweets, Following, and other Public Information"	User Metadata	Tweets you Retweet
c9	"Tweets, Following, and other Public Information"	User Metadata	"Other bits of information that result from your use of the Service"
d1	"Location Information"	User Tweet Location Info	Location in Tweets
d2	"Location Information"	User Profile Info	Your Location (profile) (appears to be the same as b2)
d3	"Location Information"	User Location Info	Trend Location
d4	"Location Information"	User Location Info	Computer Location
d5	"Location Information"	User Location Info	Mobile Location
e1	"Links"	Links clicked by User	Category: How you interact with links across our Services
e2	"Links"	Links clicked by User	How you interact with link in our email notifications
e3	"Links"	Links clicked by User	How you interact with links in third-party services
e4	"Links"	Links clicked by User	How you interact with links in client applications
f1	"Cookies"	User Cookies	Website Usage Data
f2	"Cookies"	User Cookies	Session Cookies
f3	"Cookies"	User Cookies	Persistent Cookies

code	Privacy Policy Category	General Classification of Data (Author Generated)	Detailed Type of Data
g1	"Log Data"	User Log Data	Category: Log Data
g2	"Log Data"	User Log Data	IP Address
g3	"Log Data"	User Log Data	Browser Type
g4	"Log Data"	User Log Data	Operating System
g5	"Log Data"	User Log Data	Referring Web Page
g6	"Log Data"	User Log Data	Pages Visited
g7	"Log Data"	User Log Data	Location
g8	"Log Data"	User Log Data	Mobile Carrier
g9	"Log Data"	User Log Data	Device ID
g10	"Log Data"	User Log Data	Application ID
g11	"Log Data"	User Log Data	Search Terms
g12	"Log Data"	User Log Data	Cookie Information (see category F)
h1	"Widget Data"	User Widget Data	Visits to third-party websites that integrate Twitter buttons or widgets
h2	"Widget Data"	User Widget Data	Log Data: Webpage Visited
h3	"Widget Data"	User Widget Data	Log Data: Cookie that identifies your browser
i1	"Third Parties"	Third Party Service Data about User	User Browser Information
i2	"Third Parties"	Third Party Service Data about User	User Web Page Requests
i3	"Third Parties"	Third Party Service Data about User	User Cookies
i4	"Third Parties"	Third Party Service Data about User	User IP Address

code	Privacy Policy Category	General Classification of Data (Author Generated)	Detailed Type of Data
i5	"Third Parties"	Third Party Service Data about User	User Browser Cookie ID
16	"Third Parties"	Third Party Service Data about User	Cryptograhic Hash of a common account identifier (such as e-mail)
17	"Third Parties"	Third Party Service Data about User	Google Analytics Information
i8	"Third Parties"	Third Party Service Data about User	"ads about things you may have already shown interests in" (implies interests as a form of data)

(continued on the next page)

code	When Is it Collected?	Who is it Collected From?	What is this Data Used For?
a1	Upon registration	From User	Not described
a2	Upon registration	From User	Not described
a3	Upon registration	From User	Not described
a4	Upon registration	From User	"We may use your contact information to help others find your Twitter account, including through third-party services and client applications"
b1	Upon profile setup	From User	Not described
b2	Upon profile setup	From User	Not described
b3	Upon profile setup	From User	Not described
b4	Upon profile setup	From User	Not described
b5	Upon profile setup	From User	Not described
b6	Upon profile setup	From User	Not described
b7	Upon connection with another service	From User	Not described
b8	Upon connection with another service	From Linked Services (3rd Party)	Not described
c1	During use of Twitter Services: When a user creates a tweet	From User	Not described
c2	During use of Twitter Services: When a user creates a tweet	From User	Not described
c3	During use of Twitter Services: When a user creates a tweet	Generated automatically, from user	Not described
c4	During use of Twitter Services: When a user creates a tweet	Generated automatically, from user	Not described
c5	During use of Twitter Services: When a user creates a list	Generated automatically, from user	Not described
c6	During use of Twitter Services: When a user follows another user	Generated automatically, from user	Not described
c7	During use of Twitter Services: When a user favorites a user	Generated automatically, from user	Not described

code	When Is it Collected?	Who is it Collected From?	What is this Data Used For?
c8	During use of Twitter Services: When a user retweets a tweet	Generated automatically, from user	Not described
c9	During use of Twitter Services	Generated automatically, from user	Not described
d1	During use of Twitter Services: When a user creates a tweet	From User	Not described
d2	Upon profile setup	From User	Not described
d3	During use of Twitter Services	From User	Not described
d4	During use of Twitter Services	User allows computer to indicate	Not described
d5	During use of Twitter Services	User allows mobile to indicate	Not described
e1	Upon user clicking link; collected through redirection	From User	Not described
e2	Upon user clicking link; collected through redirection	From User	Not described
e3	Upon user clicking link; collected through redirection	From User	Not described
e4	Upon user clicking link; collected through redirection	From User	Not described
f1	Upon user interaction with Services	Twitter generates / User browser accepts & retransmits where called	Not described
f2	Upon user interaction with Services	Twitter generates / User browser accepts & retransmits where called	Not described
f3	Upon user interaction with Services	Twitter generates / User browser accepts & retransmits where called	Not described
g1	"We receive Log Data when you interact with our Services, for example, when you visit our websites, sign into our Services,	User in use of Services	Not described
g2		User in use of Services	Not described

code	When Is it Collected?	Who is it Collected From?	What is this Data Used For?
g3	interact with our e-mail notifications, use your Twitter account to authenticate to a third-party website or application, or visit a third-party website that includes a Twitter button or widget"	User in use of Services	Not described
g4		User in use of Services	Not described
g5		User in use of Services	Not described
g6		User in use of Services	Not described
g7		User in use of Services	Not described
g8		User in use of Services	Not described
g9		User in use of Services	Not described
g10		User in use of Services	Not described
g11		User in use of Services	Not described
g12		User in use of Services	Not described
h1	When a user visits a 3rd party that uses Twitter buttons or widgets	From third-Parties (through user presence)	Not described
h2	When a user visits a 3rd party that uses Twitter buttons or widgets	From third-Parties (through user presence)	Not described
h3	When a user visits a 3rd party that uses Twitter buttons or widgets	From third-Parties (through user presence)	Not described
i1	"When" is not described	From third-Parties (through user presence)	Not described
i2	"When" is not described	From third-Parties (through user presence)	Not described
i3	"When" is not described	From third-Parties (through user presence)	Not described
i4	"When" is not described	From third-Parties (through user presence)	Not described
i5	"When" is not described	From third-Parties (through user presence)	Not described

code	When Is it Collected?	Who is it Collected From?	What is this Data Used For?
16	"When" is not described	From third-Parties (through user presence)	Not described
17	"When" is not described	From third-Parties (through user presence)	Not described
i8	"When" is not described	Ad-partners	Not described

(continued on the next page)

code	User Influence on Data	Public/Private Status
a1	Data is required for user to register	Listed publicly on Services
a2	Data is required for user to register	Listed publicly on Services
a3	Data is required for user to register	Not explicit in PP, assumed private.
a4	Data is required for user to register	Not explicit in PP, assumed private.
b1	Not required for use of Services	Public
b2	Not required for use of Services	Public
b3	Not required for use of Services	Public
b4	Not required for use of Services	Public
b5	Not required for use of Services	Not explicit in PP.
b6	Not required for use of Services	Not explicit in PP.
b7	Required only for linking to other services	Not explicit in PP.
b8	Required only for linking to other services	Not explicit in PP.
c1	Users decide when to tweet	Public by Default, optionally private
c2	Users decide what content to upload	Public by Default, optionally private
c3	Metadata is auto-generated upon tweet creation.	Public by Default, optionally private
c4	Metadata is auto-generated upon tweet creation.	Public by Default, optionally private
c5	Users decide to create lists	Public by Default, optionally private
c6	Users decide who to follow	Public by Default, optionally private
c7	Users decide what information to favorite	Public by Default, optionally private

code	User Influence on Data	Public/Private Status
c8	Users decide what information to retweet	Public by Default, optionally private
c9	Unknown	Public by Default, optionally private
d1	Location data is optional for tweets. User can enable for all tweets, or selective tweets.	Public by Default if shared, optionally private
d2	Location data for profile is optional, and not required for use of Services.	Public by Default if shared, optionally private
d3	Location data for profile is optional, and not required for use of Services.	Public by Default if shared, optionally private
d4	User enables computer to transmit computer location information	Public by Default if shared, optionally private
d5	User enables mobile device to communicate information	Public by Default if shared, optionally private
e1	Gathered automatically once user clicks on a link within the Services	Not explicit, however, there is a note that "We [keep track of how you interact with links] to help improve our Services, to provide more relevant advertising, and to be able to share aggregate click statistics, such as how many times a particular link was clicked on." Indicates data is shared in aggregate with advertisers.
e2	Gathered automatically once user clicks on a link within the Services	
e3	Gathered automatically once user clicks on a link within the Services	
e4	Gathered automatically once user clicks on a link within the Services	
f1	User can block cookies; might interfere with Services	Not explicit in PP, assumed private.
f2	User can block cookies; might interfere with Services	Not explicit in PP, assumed private.
f3	User can block cookies; might interfere with Services	Not explicit in PP, assumed private.
g1	Automatically generated through users interactions with Services.	Not explicit in PP, assumed private.
g2	Automatically generated through users interactions with Services.	Not explicit in PP, assumed private.
g3	Automatically generated through users interactions with Services.	Not explicit in PP, assumed private.

code	User Influence on Data	Public/Private Status
g4	Automatically generated through users interactions with Services.	Not explicit in PP, assumed private.
g5	Automatically generated through users interactions with Services.	Not explicit in PP, assumed private.
g6	Automatically generated through users interactions with Services.	Not explicit in PP, assumed private.
g7	Automatically generated through users interactions with Services.	Not explicit in PP, assumed private.
g8	Automatically generated through users interactions with Services.	Not explicit in PP, assumed private.
g9	Automatically generated through users interactions with Services.	Not explicit in PP, assumed private.
g10	Automatically generated through users interactions with Services.	Not explicit in PP, assumed private.
g11	Automatically generated through users interactions with Services.	Not explicit in PP, assumed private.
g12	Automatically generated through users interactions with Services.	Not explicit in PP, assumed private.
h1	Automatically generated through users interactions with 3rd parties that use Twitter buttons or widgets. U"this feature is optional and not yet available to all users. If you want, you can suspend it or turn it off, which removes from your browser the unique cookie that enables the feature"	Not explicit in PP, assumed private.
h2		Not explicit in PP, assumed private.
h3		Not explicit in PP, assumed private.
i1	Auto generated	Not explicit in PP, assumed private.
i2	Auto generated	Not explicit in PP, assumed private.
i3	Auto generated unless users block cookies	Not explicit in PP, assumed private.
i4	Auto generated	Not explicit in PP, assumed private.
i5	Auto generated unless users block cookies	Not explicit in PP, assumed private.
16	Auto generated	Not explicit in PP, assumed private.
17	Auto generated	Not explicit in PP, assumed private.

code	User Influence on Data	Public/Private Status
i8	Tailored ads can be turned off under privacy settings so that your account is not matched to information shared by ad partners for tailoring ads. Can also be pushed out through "Do Not Track" option, which seems to imply cookie base.	Not explicit in PP, assumed private.

(continued on the next page)

code	Noted as Shared With Anyone In Particular or Displayed in Particular Places?	Retention Notes
a1	Public, profile page, search results	No retention notes
a2	Public, profile page, search results	No retention notes
a3	Not described.	No retention notes
a4	"We may use your contact information to help others find your Twitter account, including through third-party services and client applications"	No retention notes
b1	"We may use your contact information to help others find your Twitter account, including through third-party services and client applications"	No retention notes
b2		No retention notes
b3		No retention notes
b4		No retention notes
b5		No retention notes
b6		Can be deleted
b7	Not described.	Deleted within a few weeks of your disconnected from Twitter your account on the other service.
b8	Not described.	Deleted within a few weeks of your disconnected from Twitter your account on the other service.
c1	Public by default, optionally private. Also noted, "Searchable by search engines and immediately delivered via SMS and our APIs to a wide range of users and services, with one example being the US Library of Congress, which archives tweets for historical purposes"	No retention notes in PP
c2		No retention notes in PP
c3		No retention notes in PP
c4		No retention notes in PP
c5		No retention notes in PP
c6		No retention notes in PP
c7		No retention notes in PP

code	Noted as Shared With Anyone In Particular or Displayed in Particular Places?	Retention Notes
c8		No retention notes in PP
c9		No retention notes in PP
d1	Public by Default if shared, optionally private	Location information can be deleted, however, Twitter notes (not in the PP but in the Twitter Rules) that "It is important to note that deleting location information in your settings does not guarantee the information will be removed from all copies of the data on third-party applications or in external search results."
d2	Public by Default if shared, optionally private	
d3	Public by Default if shared, optionally private	
d4	Public by Default if shared, optionally private	
d5	Public by Default if shared, optionally private	
e1	Not explicit, however, there is a note that "We [keep track of how you interact with links] to help improve our Services, to provide more relevant advertising, and to be able to share aggregate click statistics, such as how many times a particular link was clicked on." Indicates data is shared in aggregate with advertisers.	No retention notes in PP
e2		No retention notes in PP
e3		No retention notes in PP
e4		No retention notes in PP
f1	Not described.	No retention notes in PP
f2	Not described.	No retention notes in PP
f3	Not described.	No retention notes in PP
g1	Not described.	"If not already done earlier, for example, as provided below for Widget Data, we will either delete Log Data or remove any common account identifiers, such as your username, full IP address, or email address after 18 months"
g2	Not described.	
g3	Not described.	
g4	Not described.	
g5	Not described.	
g6	Not described.	

code	Noted as Shared With Anyone In Particular or Displayed in Particular Places?	Retention Notes
g7	Not described.	
g8	Not described.	
g9	Not described.	
g10	Not described.	
g11	Not described.	
g12	Not described.	
h1	Not described.	"After a maximum of 10 days, we start the process of deleting or aggregating Widget Data, which is usually instantaneous but in some cases may take up to a week."
h2	Not described.	
h3	Not described.	
i1	Not described.	No retention notes in PP
i2	Not described.	No retention notes in PP
i3	Not described.	No retention notes in PP
i4	Not described.	No retention notes in PP
i5	Not described.	No retention notes in PP
16	Not described.	No retention notes in PP
17	Not described.	No retention notes in PP
i8	Not described.	No retention notes in PP

Appendix C: User-Survey

Please note that this is only the Word version of the survey. It does not reflect the layout or skip logic used in the Qualtrics version of the survey. Questions about facets of information flow are marked as being either [correct] or [incorrect]. These statements were not visible by respondents.

Consent Block

University of Wisconsin – Milwaukee

Consent to Participate in Online Survey Research

Study Title: Informational Power on Twitter: A Mixed-Methods Exploration of User Knowledge and Technological Discourse about Information Flows

Person Responsible for Research: Nicholas Proferes, School of Information Studies, University of Wisconsin-Milwaukee

Study Description: The purpose of this research study is to better understand what people know about the popular social media platform Twitter.com. This study is open to both people who use or have previously used Twitter and people who have never used the service before. Approximately 500 subjects will participate in this study.

If you agree to participate, you will be asked to complete an online survey that will take approximately 20 minutes to complete. There are approximately 55 questions in this survey. The questions will ask demographic information about you, such as your age,

education, and your use history with Twitter. The survey will also ask you a series of questions about your familiarity with the Twitter platform.

Risks / Benefits: Risks to participants are considered minimal. Collection of data and survey responses using the internet involves the same risks that a person would encounter in everyday use of the internet, such as breach of confidentiality. While the researchers have taken every reasonable step to protect your confidentiality, there is always the possibility of interception or hacking of the data by third parties that is not under the control of the research team.

Limits to Confidentiality

Identifying information such as your e-mail address will be collected to track survey completion. Data will be retained on the Qualtrics website server for 1 year and will be deleted after this time. However, data may exist on backups or server logs beyond the timeframe of this research project. Data transferred from the survey site will be saved in an encrypted format for 7 years. Only the principal investigator (Nicholas Proferes) will have access to the data collected by this study. However, the Institutional Review Board at UW-Milwaukee or appropriate federal agencies like the Office for Human Research Protections may review this study's records. Your identifying information will be removed after the close of the survey and all study results will be reported without identifying information so that no one viewing the results will ever be able to match you with your responses.

This study was approved by the UWM IRB on Sept 30th, 2014. Its approval number is IRB #15.064.

Voluntary Participation: Your participation in this study is voluntary. You may choose to not answer any of the questions or withdraw from this study at any time without penalty. Your decision will not change any present or future relationship with the University of Wisconsin Milwaukee.

Who do I contact for questions about the study: For more information about the study or study procedures, contact Nicholas Proferes at proferes@uwm.edu.

Who do I contact for questions about my rights or complaints towards my treatment as a research subject? Contact the UWM IRB at 414-229-3173 or irbinfo@uwm.edu

Research Subject's Consent to Participate in Research:

By entering this survey, you are indicating that you have read the consent form, you are age 18 or older and that you voluntarily agree to participate in this research study.

Thank you!

Q1. Please indicate your willingness to participate by selecting one of the following:

□ I agree to take part in this survey.

□ I do not agree to take part in this survey.

Demographics

Q2. Please indicate your age range:

□ 18-24

□ 25-29

□ 30-39

□ 40-49

□ 50-59

□ 60 or above

Q3. Please indicate your gender:

_____

Q4. Please indicate the highest level of education you have completed:

□ Some High School Education

□ Finished High School Degree

□ Some Undergraduate Education

□ Finished Undergraduate Degree

□ Some Graduate-Level Education

□ Finished Graduate or Other Post-Undergraduate Professional Degree

Q5. Have you ever used the social media platform Twitter?

□ No, I have never used Twitter.

□ I have visited the Twitter website before (Twitter.com), but do not have a registered account.

□ I have a registered account on Twitter.

□ I previously had an account on Twitter, but deleted it.

Twitter Use

Q5a. How do you access Twitter? (Please select all that apply)

□ I use the Twitter.com website via my web-browser to access Twitter.

□ I use a desktop application, such as TweetDeck, to access Twitter.

□ I use a mobile application, such as the Twitter app, to access Twitter.

□ I use some other means to access Twitter (please specify):

Q5b. Which way of accessing Twitter would you say you use the most often?

□ I use the Twitter.com website via my web-browser most often to access Twitter.

□ I use a desktop application, such as TweetDeck, most often to access Twitter.

□ I use a mobile application, such as the Twitter app, most often to access Twitter.

□ I use some other means to access Twitter most often (please specify):

Q5c. When was the last time you posted a Tweet?

□ I have never sent a tweet.

□ The last time I sent a tweet was over a year ago.

□ The last time I sent a tweet was over a month ago.

□ The last time I sent a tweet was over a week ago.

□ The last time I sent a tweet was earlier this week.

Q5d. When was the last time you visited the Twitter.com homepage?

□ The last time I visited the Twitter.com was over a year ago.

□ The last time I visited the Twitter.com was over a month ago.

□ The last time I visited the Twitter.com was over a week ago.

□ The last time I visited the Twitter.com was earlier this week.

Q5e. I would say that I access Twitter:

□ Almost never.

□ Occasionally.

□ Fairly Regularly.

□ Often.

Q5f. Roughly how long ago did you first sign up for Twitter?

_____

Data Questions

For the following questions, we will be asking you a series of questions about Twitter. Please select what you believe to be the correct answer.

Please do not use any external sources to help you answer these questions. You may choose "I don't know the answer to this question" if you feel that you do not know the answer.

Q6. Messages on Twitter (also called ‘Tweets’) are limited to 210-characters in length. [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q7. Users can select an option when composing a tweet to share location information, such as their GPS coordinates, along with their tweet. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q8. Twitter does not ever withhold tweets or user accounts from being accessed within specific countries, even if they have received a legal request to do so. [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q9. If you have a “protected” account on Twitter, your tweets are only visible to the users that follow you and to the users that follow your followers. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q10. Twitter warns users if a link embedded in someone else's tweet has been marked as "possibly sensitive." [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Protocol Questions

For the following questions, we will be asking you a series of questions about Twitter. Please select what you believe to be the correct answer.

Please do not use any external sources to help you answer these questions. You may choose "I don't know the answer to this question" if you feel that you do not know the answer.

Q11. Including a hashtag (the '#' symbol) in front of a word is a way of marking keywords or topics in a tweet and is sometimes used by users as a way to categorize messages. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q12. Including the "@" symbol and another user's Twitter username (such as "@PBS") at the beginning of a tweet means that Twitter will treat that as a "reply" to that user. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q13. If you "favorite" another user's tweet on Twitter, that information is only shared with the person who created the tweet. [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q14. Following someone on Twitter means that you are subscribing to their Tweets as a follower and their updates will appear in your Home tab. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Algorithm Questions

For the following questions, we will be asking you a series of questions about Twitter. Please select what you believe to be the correct answer.

Please do not use any external sources to help you answer these questions. You may choose "I don't know the answer to this question" if you feel that you do not know the answer.

Q15. Twitter user “@Jane” has a protected account. @Jane sends a @reply to Twitter user “@PBS.” @PBS will still be able to see @Jane’s tweet, even if they are not following her. [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q16. Twitter's trending topic algorithm identifies topics that are immediately popular, rather than topics that have been popular for a while or on a daily basis. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q17. All users see the same trending topics. [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q18. A Twitter user will only see another user's @replies in their home timeline if they are following both the sender and recipient of the @reply. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q19. If you are using Twitter and someone sends you a @reply, but you are not following the user, the reply will still appear on your Tweets timeline. [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Defaults Questions

For the following questions, we will be asking you a series of questions about Twitter. Please select what you believe to be the correct answer.

Please do not use any external sources to help you answer these questions. You may choose "I don't know the answer to this question" if you feel that you do not know the answer.

Q20. By default, users can receive Direct Messages from any other users. [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q21. By default, your tweets are “protected,” meaning that, unless you change a setting, your tweets will only be accessible to your followers. [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q22. Unless you make changes to the default choices on your Twitter settings page, Twitter tailors its suggestions of the people you might enjoy following based on your recent visits to other websites that have integrated Twitter buttons or widgets. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q23. Unless you make changes to the default choices on your Twitter settings page, location information (such as GPS coordinates) about where you are tweeting from is automatically made publicly accessible along with your tweets. [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q24. Unless you make changes to the default choices on your Twitter settings page, Twitter tailors the advertisements you receive based on the information it gets about you from third-parties. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Interface Questions

For the following questions, we will be asking you a series of questions about Twitter. Please select what you believe to be the correct answer.

Please do not use any external sources to help you answer these questions. You may choose "I don't know the answer to this question" if you feel that you do not know the answer.

Q25. Twitter offers interfaces for programmers and application developers called the Twitter APIs, which can be used to make applications, websites, widgets, and other projects that interact with Twitter. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q26. Old tweets are automatically deleted from Twitter’s servers after 2 years. [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q27. When you visit a website with Twitter buttons or widgets like the “Tweet This” button, Twitter does not receive information about that visit unless you click on the button or widget. [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q28. Twitter offers a search interface to programmers that allows them to search for tweets by GPS data (latitude, longitude and radius area), and will attempt to find tweets created by users whose profile location matches the search parameters. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q29. There is no way for a user to retrieve all of the tweets they have ever created. [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q30. When you visit a website with Twitter buttons or widgets like the “Tweet This” button, Twitter may receive information about the URL (web-address) of the website you are visiting. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Business Model

For the following questions, we will be asking you a series of questions about Twitter. Please select what you believe to be the correct answer.

Please do not use any external sources to help you answer these questions. You may choose "I don't know the answer to this question" if you feel that you do not know the answer.

Q31. Please identify which of the following are ways Twitter that generates revenue (you can make multiple selections for this question):

□ Twitter generates revenue by selling advertising in the form of "promoted tweets" that appear in users’ timelines. [correct]

□ Twitter generates revenue by charging people for verified accounts. [incorrect]

□ Twitter generates revenue by selling access to the full stream of real-time tweets created by users to third-parties. [correct]

□ Twitter generates revenue by charging other websites to put Twitter buttons and widgets on their sites. [incorrect]

□ Twitter generates revenue by selling advertising in the form of "promoted trends" that appear in in the “Trending Now” section of the site. [correct]

□ Twitter generates revenue by charging news outlets like CNN every time they talk about things that have happened on Twitter. [incorrect]

□ Twitter generates revenue by selling advertising in the form of "promoted accounts" that are recommended to users. [correct]

□ Twitter generates revenue by charging websites that have Twitter buttons every time a visitor to those websites clicks on a “Tweet This” button. [incorrect]

□ I don’t know the answer to this question.

Governance

For the following questions, we will be asking you a series of questions about Twitter. Please select what you believe to be the correct answer.

Please do not use any external sources to help you answer these questions. You may choose "I don't know the answer to this question" if you feel that you do not know the answer.

Q32. Twitter has three documents that govern users on the site, the “Terms of Service,” the “Privacy Policy” and “The Twitter Rules.” [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q33. Twitter allows spam, abuse, phishing, and malware on its platform as long as it is marked as “potentially sensitive.” [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q34. Users of Twitter's APIs are forbidden from selling, renting, leasing, or redistributing access to the Twitter API or Twitter Content to any third party without prior written approval from Twitter. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q35. The purpose of this question is to assess your attentiveness to question wording. For this question please mark the "I don’t know the answer to this question" response.

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q36. Twitter advertisers are prohibited from offering adult or sexual products or services, drugs or drug paraphernalia, endangered species products, hate content, tobacco, unauthorized ticket sales, or weapons. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q37. All tweets created by users with public accounts are given to the Library of Congress for archival by Twitter. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Users

For the following questions, we will be asking you a series of questions about Twitter. Please select what you believe to be the correct answer.

Please do not use any external sources to help you answer these questions. You may choose "I don't know the answer to this question" if you feel that you do not know the answer.

Q38. Unregistered visitors to Twitter.com can still view publicly created tweets but cannot use the "search" feature of the website. [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q39. "Verified accounts" on Twitter are those for which the user has paid Twitter. [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q40. Twitter's "Certified Products" partners are all prohibited from reselling historical Twitter data, such as old tweets, to third-parties. [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q41. If you have a “protected” account on Twitter and another user wants to follow you, you must approve them before they can do so. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q42. "Verified accounts" on Twitter are those for which the identities of the individuals or brands they represent have been authenticated. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q43. Advertisers who purchase "promoted tweets" on Twitter receive information about the number of users that have been served that tweet, the number of clicks on a promoted tweet (and information on which piece of the tweet users clicked on), the number of times a tweet has been retweeted, and the number of times a promoted tweet has been replied to. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q44. Information about the number of tweets, number of photos, number of followers and followees, and number of favorites of users with protected accounts is not publicly accessible information. [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Ownership

For the following questions, we will be asking you a series of questions about Twitter. Please select what you believe to be the correct answer.

Please do not use any external sources to help you answer these questions. You may choose "I don't know the answer to this question" if you feel that you do not know the answer.

Q45. Twitter is a publicly traded company on the New York Stock Exchange. [correct]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Q46. Tim Cook is the current CEO of Twitter. [incorrect]

□ Yes, this is correct.

□ No, this is incorrect.

□ I don't know the answer to this question.

Habits

Q47. Pick the option that would best describe your opinion: I feel like Twitter is discussed in the news…

□ Never.

□ Occasionally.

□ Sometimes.

□ Regularly.

Q48. Pick the option that would best describe you: I keep up with news about Twitter…

□ Never.

□ Occasionally.

□ Sometimes.

□ Regularly.

Q49. Users of Twitter who have registered accounts have to agree to the Terms of Service when they sign up for the service. Select the option below that best describes how closely you would say you have read the Terms of Service.

□ I have never read the Terms of Service.

□ I have skimmed over the Terms of Service.

□ I have read the Terms of Service in some detail, but not fully.

□ I have fully read the Terms of Service in detail.

Q50. Twitter offers a Privacy Policy to users that discusses how it collects, stores and shares user information. Select the option below that best describes how closely you would say you have read the Privacy Policy.

□ I have never read the Privacy Policy.

□ I have skimmed over the Privacy Policy.

□ I have read the Privacy Policy in some detail, but not fully.

□ I have fully read the Privacy Policy in detail.

Q51. Twitter has a set of "Twitter Rules" that discusses how it handles certain types of content on Twitter. Select the option below that best describes how closely you would say you have read the Twitter Rules.

□ I have never read the Twitter Rules.

□ I have skimmed over the Twitter Rules.

□ I have read the Twitter Rules in some detail, but not fully.

□ I have fully read the Twitter Rules in detail.

Appendix D: Recruitment E-mail for Survey

From: Nicholas J Proferes <noreply@qemailserver.com>

Subject: Research Study on Perceptions of Twitter

Good Morning,

My name is Nicholas Proferes and I am a doctoral student in the School of Information Studies at UWM. I am currently investigating public perceptions of the social media platform Twitter as part of my dissertation research and I am e-mailing you to invite you to participate in this study. The survey is open to all and you do not need to be a user of Twitter in order to participate. Please read the details below and then click on the following link to take the brief online survey. Thank you, in advance, for your contribution!

Follow this link to the Survey:

Take the Survey

Or copy and paste the URL below into your internet browser:

[URL]

Study Title: Informational Power on Twitter: A Mixed-Methods Exploration of User Knowledge and Technological Discourse about Information Flows

Principal Investigators: Nicholas Proferes

Study Purpose and Rationale

The purpose of this research study is to better understand what people know about the popular social media platform Twitter.com. Approximately 500 subjects will participate in this study. If you agree to participate, you will be asked to complete an online survey that will take approximately 20 minutes to complete. The questions will ask demographic information about you, such as your age, gender, education, and your use history with Twitter. It will also ask you a series of questions about how the Twitter platform works.

Inclusion/Exclusion Criteria

To participate in this study you must be 18 years or older. You do not need to be a user of Twitter in order to take the survey, and the survey is open to students, staff and faculty.

Data Confidentiality

Identifying information such as your e-mail address will be collected to track survey completion. Data will be retained on the Qualtrics website server for 1 year and will be deleted after this time. However, data may exist on backups or server logs beyond the timeframe of this research project. Data transferred from the survey site will be saved in an encrypted format for 7 years. Only the principal investigator (Nicholas Proferes) will have access to the data collected by this study. However, the Institutional Review Board at UW-Milwaukee or appropriate federal agencies like the Office for Human Research Protections may review this study's records. The researcher will remove your identifying information after the survey closes and all study results will be

reported without identifying information so that no one viewing the results will ever be able to match you with your responses.

Risks / Benefits: Risks to participants are considered minimal. Collection of data and survey responses using the internet involves the same risks that a person would encounter in everyday use of the internet, such as breach of confidentiality. While the researchers have taken every reasonable step to protect your confidentiality, there is always the possibility of interception or hacking of the data by third parties that is not under the control of the research team.

IRB Contact Information

This study was approved by the UWM IRB on Sept 30th, 2014. Its approval number is IRB #15.064. For questions about your rights as a research subject Contact the Director, University of Wisconsin- Milwaukee, IRB at 414-229-3173 or irbinfo@uwm.edu.
For questions about your rights as a research subject, please contact Director, Office of Research Integrity, Ball State University, Muncie, IN 47306, (765) 285-5070, irb@bsu.edu.

Researcher Contact Information

Principal Investigators:

Nicholas Proferes

Ph.D. Candidate, School of Information Studies

University of Wisconsin-Milwaukee

Milwaukee, WI 53211

proferes@uwm.edu

Follow the link to opt out of future emails:

Click here to unsubscribe

Curriculum Vitae

Nicholas J Proferes

proferes@uwm.edu

Ph.D. Candidate, School of Information Studies, University of Wisconsin-Milwaukee

EDUCATION

Ph.D. Candidate	**University of Wisconsin - Milwaukee** Information Studies, 2010 – present Expected Graduation Date: May 2015 Dissertation Title: "Informational Power on Twitter: A Mixed-Methods Exploration of User Knowledge and Technological Discourse about Information Flows" Chair: Dr. Michael Zimmer
M.A.	**Georgetown University** Communication, Culture and Technology, May 2008 Thesis: "Privacy: How Do We Define It, Assess It, and Seek to Protect It Online" Committee: Dr. Diana Owen (chair), Rev. Dr. Eric Zimmer
B.S.	**George Mason University** Information Technology, Spring 2006 Major Area: Network Administration and Security

PROFESSIONAL POSITIONS HELD

2008-2010	**National Science Foundation - Arlington, VA.** Science Assistant - Office of Integrative Activities (OIA)
2008	**Internet Education Foundation – Washington, D.C.** Google Policy Fellow
2007	**International Resources Group – Washington, D.C.** Knowledge Management Specialist

PUBLICATIONS

Proferes, N. (2015). Legal Limits To Technology Use. In R. Mansell & P. Ang (eds.), *International Encyclopedia of Digital Communication & Society*. Hoboken, NJ: Wiley.

Zimmer, M. & **Proferes, N.** (2014). A Topology of Twitter Research: Disciplines, Methods, and Ethics. *Aslib Journal of Information Management* 66(3), 250-261.

(continued on next page)

Proferes, N. (2014). What Happens to Tweets? Descriptions of Temporality in Twitter's Organizational Rhetoric. *iConference Proceedings*. Berlin, Germany.

Zimmer, M. & **Proferes, N.** (2013). Privacy on Twitter / Twitter on Privacy. In K. Weller, A. Bruns, J. Burgess, M. Mahrt & C. Puschmann (eds.), *Twitter and Society*. New York, NY: Peter Lang.

Proferes, N. (2012). [Book Review] Delete: The Virtue of Forgetting in the Digital Age. Viktor Mayer-Schönberger. Princeton, NJ: Princeton University Press, 2009. *The Journal of Popular Culture*, 45(1), 226–228. doi:10.1111/j.1540-5931.2011.921_1.x

Adamick, J., Buchanan, E., Fountain, J., Goncalves, M.S., **Proferes, N.** (2010). [White Paper] Advancing Ethical Research Across Disciplines, NSF SBE 2020. http://www.nsf.gov/sbe/sbe_2020/2020_pdfs/Adamick_Jessica_139.pdf

Proferes, N. (2007). Modern Surveillance Methods and Public Trust. *Gnovis*, 8(1), 75-82.

PRESENTATIONS

2014 *What Happens to Tweets? Descriptions of Temporality in Twitter's Organizational Rhetoric*. iConference 14: Berlin, Germany.

2013 *Users, Myth, and Social Media [Student Presentation]*. Oxford Internet Institute Summer Doctoral Program: *Toronto, ON.*

2012 *The Ethics of Twitter Research: A Topology of Disciplines, Methods and Ethics Review Boards*
Co-presented with Dr. Michael Zimmer
2nd International Symposium on Digital Ethics: *Chicago, IL*.

2012 *The Political Economy of Reputation Online: Karma, Reddit, and the Knowledge Factory*
Midwest Interdisciplinary Graduate Conference: *Milwaukee, WI.*

2011 *Oh, the Ethics You'll Know [IGNITE Presentation]*
Association of Internet Researchers 12: Seattle, WA.

POSTER PRESENTATIONS

2015 *An Examination of Library and Information Studies Faculty Experience with and Attitudes toward Open Access Scholarly Publishing*. iConference 2015: Newport Beach, CA.

2015 *Exploring User Knowledge of Information Flows on Twitter*. ALISE: Chicago, IL.

RESEARCH AND TEACHING ACTIVITIES

Fall 2014, Spring 2015	**Project Assistant** University of Wisconsin-Milwaukee Center for 21st Century Studies
Spring 2014	**Instructor** University of Wisconsin-Milwaukee INFOST 120 – Information Technology Ethics
Spring 2013, Fall 2013	**Instructor** University of Wisconsin-Milwaukee INFOST 110 – Introduction to Information Science
Spring 2013	**Editorial Assistant** University of Wisconsin-Milwaukee Law & Policy Volume of the International Encyclopedia of Digital Communication & Society; Editor: Dr. Sandra Braman
Fall 2012, Summer 2013	**Research Assistant** University of Wisconsin-Milwaukee Dr. Sandra Braman
Fall 2012	**Teaching Assistant** University of Wisconsin-Milwaukee L&I SCI 110 – Introduction to Information Science Instructor: Dr. Michael Zimmer
Summer 2012, Summer 2013, Summer 2014	**Project Assistant** University of Wisconsin-Milwaukee Dr. Michael Zimmer
Fall 2012, Summer 2013, Summer 2014	**Project Assistant** University of Wisconsin-Milwaukee Dr. Wilhelm Peekhaus
Fall 2010 - Spring 2012	**Project Assistant** UWM Responsible Conduct of Research (RCR) Training Program University of Wisconsin-Milwaukee Dr. Elizabeth Buchanan, Dr. Janis Eells, Dr. Marjorie Piechowski

FELLOWSHIPS, AWARDS AND GRANTS

2014	**Chancellor's Graduate Student Award** School of Information Studies, University of Wisconsin – Milwaukee
2014	**Graduate Teaching Award** School of Information Studies, University of Wisconsin – Milwaukee
2013	**Chancellor's Graduate Student Award** School of Information Studies, University of Wisconsin – Milwaukee

2013 **Graduate Teaching Award**
School of Information Studies, University of Wisconsin – Milwaukee

2013 **Travel Award**
Graduate School, University of Wisconsin – Milwaukee

2013 **Doctoral Research Grant**
School of Information Studies, University of Wisconsin-Milwaukee

2013 **Featured Graduate Student**
Graduate School, University of Wisconsin – Milwaukee

2012 **Chancellor's Graduate Student Award**
School of Information Studies, University of Wisconsin – Milwaukee

2011 **Chancellor's Graduate Student Award**
School of Information Studies, University of Wisconsin – Milwaukee

2011 **Travel Award**
Graduate School, University of Wisconsin – Milwaukee

PROFESSIONAL SERVICE

2014 Reviewer for *New Media & Society*.

2013 President, School of Information Studies Doctoral Student Organization.

2013 PhD Student Representative, School of Information Studies Doctoral Program Committee.

2013 Editor, School of Information Studies Newsletter.

2012 PhD Student Representative, UWM Learning and Technology Center Review Committee.

2012 PhD Student Representative, UWM Responsible Conduct of Research Advisory Group.

2011 Reviewer for Great Lakes Graduate Conference

2010 - 2011 PhD Student Representative, School of Information Studies Diversity and Equity Committee.

2010 - 2015 Judge for UWM Undergraduate Research Fair

PROFESSIONAL AFFILIATIONS

2013 - Current American Association for the Advancement of Science

2012 - Current Social Studies of Information Research Group

2010 - Current Association of Internet Researchers