American Behavioral Scientist
http://abs.sagepub.com/

Connecting Theory to Social Technology Platforms: A Framework for Measuring Influence in Context

Sean Goggins and Eva Petakovic

American Behavioral Scientist published online 7 April 2014
DOI: 10.1177/0002764214527093

The online version of this article can be found at:
http://abs.sagepub.com/content/early/2014/04/07/0002764214527093

Published by:
SAGE
http://www.sagepublications.com

Article

American Behavioral Scientist
1–17

DOI: 10.1177/0002764214527093
abs.sagepub.com

Connecting Theory to Social Technology Platforms: A Framework for Measuring Influence in Context

Sean Goggins[1] and Eva Petakovic[1]

Abstract

In this article, the authors synthesize 3 years of social technologies research, including studies of Facebook, Twitter, and GitHub, to present a theory driven framework to guide future social scientific research using "Big Data." They connect levels of analysis derived from empirical study of influence to the electronic trace data generated by social technologies. Specifically, the authors outline a relationship between social media technology platforms, individual goals for participation, and emergent small groups to inform future research on influence in social technologies. They incorporate theory from small group literature, communities and networks of practice, and media theory to explicate a contextual framework for measuring influence. In their discussion, the authors build on the contrast between influence indicators in Facebook, Twitter, and GitHub to argue for a greater focus on the influence abstractions of articulation and affiliation.

Keywords

influence, levelism, electronic trace data, theory

Introduction

Influencing children to eat their vegetables and influencing colleagues to allocate resources for a new lab require different strategies. Specific power relations and resource requirements exist in each case, and the way that we address these elements differs when influencing an individual, a group, an organization, or a society. Most of us have some experience influencing or failing to influence people in these types of

[1]University of Missouri, Columbia, MO, USA

Corresponding Author:
Sean Goggins, University of Missouri, Columbia, MO, USA.
Email: s@goggins.com

relationships. Our experience of influencing using social technology within and across social groupings is limited, and consequently, the functions of influence in social technologies are little understood. Given the growing presence of social technologies in our lives, understanding influence in these contexts is valuable.

Influence is not a singular construct. From one perspective, influence is a complex social process leading to adoption of new behaviors by groups of people (Malhotra & Galletta, 2003). Effective strategies for advancing uniform social behaviors include sanctioning deviants (punishment), positive payoff externalities (reward), conformity preferences (belonging), and communication between people (rhetoric). The first three strategies address reward structures and are backed by extensive cross-disciplinary social science research. Examining influence from the perspective of communication, in contrast, implies that convergence toward a particular behavior will occur if communication is credible and costless (Bikhchandani, Hirshleifer, & Welch, 1992). Communication through social technologies creates a new path for influence that alters the landscape of human relations.

Finer parsing of social technology practice, reach, and effect is necessary for understanding and measuring influence across social technology platforms (Goggins & Mascaro, 2011). The relatively low cost of engagement through social technologies is well documented (De Souza & Preece, 2004), but how engagement translates to influence in social technologies is not well understood (King, 2011).

A number of studies note that engagement and participation in social technology are both constrained and enabled by the technology itself (Gilbert & Karahalios, 2009; Grabowicz, Ramasco, & Eguiluz, 2012; Wellman, Haase, Witte, & Hampton, 2001). Facebook reinforces relationships in the physical world, Twitter embodies a key source of contemporary information diffusion enabling loose, topical connections between individuals (Golder & Yardi, 2010; Yardi & boyd, 2010), and GitHub is a site for highly fluid, distributed, and socially engaging work (Dabbish, Stuart, Tsay, & Herbsleb, 2012). Social technology is therefore not a single construct but a rapidly changing collage of technologies and practices.

Most empirical results are shaped by the properties of one specific platform being studied. Contrasts and synthesis across platforms are limited. Furthermore, many site-specific studies either do not shape their inquiry theoretically or rely on theory grounded in studies of social phenomena from the physical world. Other studies rely heavily on computational approaches with limited backing from social scientific theory (Goggins, Mascaro, & Valetto, 2013). Such dissonance creates challenges for social scientists examining influence in social technologies (Urry, 2000).

Two specific, interrelated challenges emerge from this dissonance. First, it is largely recognized that online communities influence and are influenced by offline communities (Valenzuela, Park, & Kee, 2009), however, little work focuses on framing influence differences and interactions across social technologies. Second, the study of social phenomena on large-scale social networking sites (SNSs) like Facebook and Twitter is limited in the explanation and development of theories to explain causal connections between behaviors in social technology and socially valuable and measureable outcomes (Steinfield, Ellison, Lampe, & Vitak, 2012). Measuring influence in social technologies therefore requires that studies of influence on specific social technology platforms incorporate a systematic methodological approach and both leverage and advance coherent theories that help to explain how people engage with a particular platform (Goggins et al., 2013; Howison, Wiggins, & Crowston, 2012). Where relevant theories do not guide us sufficiently, qualitative and comparative case studies focused on measuring influence within, and more importantly across, specific social technology platforms are appropriate (Eisenhardt, 1989).

In this article, we describe common considerations for examining individual social technology platforms as a challenge of Big Social Data (BSD), arguing that large sets of social interaction data require firmer grounding in social theories than is presently established. Next, we synthesize literature related to influence within each social technology platform, creating a matrix of theoretical considerations for understanding influence across platforms. We then distill our research on influence in Facebook, Twitter, and GitHub, drawing on our empirical work related to influence on each platform. Finally, we draw on the philosophical train of thought called levelism, which makes the relationship between epistemological and ontological abstraction explicit and supports coherent theory development related to influence in social technologies.

Literature

Big Social Data

Understanding influence in social technologies relies on analysis of data generated through interactions between people. "Big Data" is a common idiom for BSD, but it is a term whose meaning is inconsistently understood. The more specific term, BSD, lays the groundwork for understanding social processes operating on a worldwide scale. The potential for such advances is great due to the ubiquity of social technologies. Interactions can be of many kinds (communication, transaction, reaction, relationship) and observed at the level of individuals, groups and organizations, and nations. When people interact through the web, mobile devices, and distributed sensors, digital traces of these interactions are left behind. These interactions are more easily quantified through digitization and sharing of document and image archives. Consequently, we face a deluge of data from which new scientific, economic, and social value can be extracted.

Nowhere are transformational opportunities and scientific challenges greater than in the social sciences. Lazer (2009) argued that the "digital breadcrumbs" of contemporary life offer "the potential of transforming our understanding of our lives, organizations, and societies in a fashion that was barely conceivable just a few years ago" (p. 2). Watts (2007) asserted that "social science is the science of the 21st century." There is a gap between existing social science theories and the new methods necessary to realize the transformative potential of these digital breadcrumbs. Lazer (2011) warned that social scientists are not equipped with methods, algorithms, or access to computing infrastructure for collecting and managing BSD. Consequently, a large portion of this data-intensive social science occurs but in places like Google, Yahoo!, and the National

Security Agency, where computational approaches capable of processing data on a grand scale have been developed. These approaches are frequently applied in ways that render them of little value for unearthing new insights into people and how they work. Boyd (2010) noted that "many . . . are approaching Big Data from a computational perspective, but it's absolutely crucial that you understand that you're dealing with data about people." What we need to do is connect social scientists, who have theoretical frameworks for studying people, with computer scientists, who have knowledge of advanced, contemporary computational methods for data analysis on a grand scale.

Big Social Data generated through participation on sites such as Facebook, Twitter, and GitHub poses a number of scientific validity and collaboration challenges for social scientists (Miller, 2011). First, there are challenges with respect to control of random variance, especially because many important mediating variables, such as age, gender, socioeconomic status, ethnicity, and political affiliation of those involved in the interactions, are unknown. Yet, many times, such variables act as important mediators in the kinds of effects that are of interest to social scientists. This poses problems in terms of interpretability, particularly because it cannot be assumed that a random sample of data has representative distribution of such variables. Situating specific findings in the context of existing theory and prior studies is therefore difficult (Goggins et al., 2013).

Influence and Social Technologies

Researchers do not treat organization studies, small group research, or information diffusion as phenomena explainable by a grand theory of social influence. Instead, social scientists recognize that specific social contexts require theories and methods of study addressing the particulars of those contexts. Studying influence in social technologies must then take into account the particulars of each social and technical context in order to operationalize influence and identify context-specific mechanisms for measuring influence.

Each social technology is used differently and for varying purposes; these differences have corollaries in theories of influence in the physical world. In this section, we consider typical use scenarios for each social technology and explicate theoretical foundations for the study of influence in specific sociotechnical contexts. As a framework for building understanding of influence in sociotechnical context, we begin by explicating forms of social connection and the nature of communication instances on Twitter, Facebook, and GitHub. We then frame these differences as composed of different epistemic, ontological levels.

Facebook. The SNS Facebook relies on mutual relations for connection. Contributions and comments center on an individual user or organization page. The context is an identifiable person or group, and the whole experience centers on the relationship between a user and his or her connections. Communication is casual and has no text limit. People use Facebook to make or maintain (Lampe, Ellison, & Steinfield, 2006) social connections with people they already know.

When researchers study Facebook, trace data are returned in the form of a complex graph of people, header posts on their pages, and responding comments (de Zúñiga, Jung, & Valenzuela, 2012).

Twitter. Twitter connections are directional; posts can be seen and followed without reciprocation. Interaction is focused more on the tweets themselves, which emerge from topical networks driven by hashtags or from follower relations indicated by user names. Furthermore, user experience on Twitter is contingent on the particular interface that a user participates through, such as Hootsuite and TweetDeck, which affects their social and navigational behavior.

Twitter communication includes three signaling markers central to information and social connections. We follow a topic through the use of a hashtag (#); we follow a person by referencing his or her username, prefixed with "@." In a small percentage of tweets, users direct message or refer to other users whom they do not need to be "connected to." Where Facebook has a "like" function for posts and comments, Twitter has a "favorite this" and the ability to "retweet" specific tweets to followers. The text of the tweet and a user's choice of signaling markers constitute a nuanced and user-controlled set of information diffusion practices. Twitter is in many respects a participatory form of mass media. User practice suggests that the context (Dourish, 2004) created on Twitter is socially connected and information diffusion focused (Chang, 2010).

GitHub. Socially connected computing has its roots in the study of corporate portals and groupware (Bansler & Havn, 2006). GitHub is a more public, more social form of contemporary groupware better optimized for software work (Dabbish et al., 2012). Unlike the traditional tools used to manage distributed software work, GitHub integrates the work of changing code, discussion of software issues, and enabling of social connection through a socially translucent (Erickson & Kellogg, 2000) user interface. GitHub is a social site based on the distributed software configuration management in Git (Bird et al., 2009) and, unlike comparable tools, does not rely on centralized control. Instead, GitHub lets any user submit candidate changes to a repository through a controlled "pull request," initiating discussion within the project (Finley, 2011). Following discussion, a project leader chooses whether to commit the proposed changes to the main repository.

Analysis of GitHub participation incorporates a more elaborate and multivalent graph of participation traces than Facebook. People come to GitHub to build professional credibility or to make a contribution to a project they rely on in their own work. GitHub users participate through discussion of issues, code contributions, and comments on pull requests. Influence, therefore, can occur through the contribution of work (code), negotiation processes around code (issue discussion), or the identification of a candidate change to the code (issue). Like any project, there are roles to be earned, and roles on GitHub are awarded based on influence, operationalized explicitly as acknowledgment of a participant's quality of work and amount of contribution.

Contrasting influence constructs across social technologies. Social technologies are primarily sites for communication-style social influence (Malhotra & Galletta, 2003). Fewer mechanisms exist for extrinsic incentive, punitive measures, or creating a sense of belonging. It is not that these other types of influence do not exist in specific ways. Bridging and bonding social capital are widely discussed, for example (Yuan & Gay, 2006). However, our argument is that communication-based influence—both the use of rhetoric and information sharing—is prominent within social technology practice and readily visible in trace data. The fact that behavioral traces are readily available in each social technology enables research focused on communication-based influence. In particular, we are able to observe how both the practices of influence and the most influential individuals change over time as a result.

Differences in the practices and purposes of participation in each of these social technologies can be conceptualized in terms of the levels of data we observe (ontology) and levels of phenomena we try to explain (epistemology). Abstractions of data and explanation are a foundation of how we think about the differences across both present and future social technologies.

Levelism and influence in social technologies. Levelism is a branch of philosophy concerned with how levels of abstraction (LoAs) are applied in scientific inquiry. Main categories of abstraction in Floridi's (2008) method of levels of abstraction include levels of organization (LoOs), which make ontological commitments, and levels of explanation (LoEs), which make epistemological commitments. The construct of gradients of abstraction (GoAs) is applied to suggest levels of greater or lesser abstraction, with all components of the higher levels composed of lower levels. Floridi's (2008) method of levelism foregrounds the connection between how we organize data and the phenomena we are attempting to explain.

Our purpose in introducing the philosophical notion of levelism is to point out that making social sense of discrete data is a problem long reasoned about abstractly in the foundational scientific field of philosophy (Floridi, 2008). The challenge of reasoning about discrete electronic trace data is addressed here as a special case of this more general problem in science. Our explanations of the patterns of influence in social technologies can be placed on firmer ground through systematic reasoning about the relationship between aggregations of observable trace data. Specifically, we are looking to avoid making hidden epistemological commitments about how influence can be observed through greater clarity in our choice of ontological commitments (i.e., the organization and aggregation of trace data). Making clear ontological and epistemological commitments in the analysis of trace data from social technologies is a distinctly sociotechnical extension of prior studies of society and culture.

Context-Sensitive Measurements of Influence

Participation and affiliation: operationalizing levelism. In this section, we outline abstractions of participation and affiliation to discover measures of influence in social technology trace data. Affiliation is expressed in a way that is specific to each social

Table 1. Examples of Participation and Affiliation Across Social Technologies.

Technology	Participation	Affiliation
Facebook	commenting on another person's post	friending a person
	liking another person's post	liking a page
	tagging a person in a photo	joining a group
Twitter	retweeting	following a person
	reply-to	
	mention	
		following a hashtag
	using a hashtag	
	favoriting a tweet	
GitHub	forking a project	following a project
	issuing a pull request	following a contributor
	commenting on a pull request	
	submitting an issue	

technology. Participation is multivalent within and across social technologies and closely tied to motives for participation. In Table 1, we identify the participation and affiliation practices that inform measures of influence in each social technology.

In this section, we discuss traces of interaction as concrete data used to explain influencing behaviors, build models, and develop theory in social technologies. Operationalizing influence through measurable acts of participation and affiliation will require us to explain levelism in the context of communication-oriented influence in social technologies (see Floridi, 2008, for a thorough explication of a method of levelism). Influence is what we are trying to explain, which is organized by LoEs. In traditional philosophy, this is epistemology.

To explain influence (our LoE) across social technologies, we first need to be specific about the levels of organization, which represent ontological commitments in traditional philosophy. A level of organization helps to conceptualize our organization of electronic trace data, which is operationalized at different levels as practices and purposes for participation in each social technology. To measure and instrument influence across technologies, we will specify cumulative pyramids of LoOs for each social technology and describe the shared abstractions of participation and affiliation across technologies.

Participation and affiliation behaviors across social technologies and within social technologies are practices through which influence occurs. To operationalize influence in each social technology, we must identify and explore the LoO-specific participation and affiliation in each social technology. The critical difference between a LoO and *conceptual schemes*, which organize experience from the perspective of participants, is separating the analytical constructs (LoO [ontology] & LoE [epistemology]) from participant experience. In the following sections, we clarify the relationship between these abstractions in advance of the analysis.

Informing Weighted, Network Analytic Analysis of Influence

Our analysis builds on empirical work that we have published related to influence on Facebook, Twitter, and GitHub. This work involves weighted social network analysis built on the explicated LoOs. We use measures of network degree centrality for participants and reveal influence changes over time through sociograms. We apply both degree centrality, which is a social network measure of direct influence, and betweenness centrality, which is a social network measure of brokerage influence, as measures in our empirical work. Our approach differs from much work that uses network centrality measures for actors. We are explicit about the relationship between measures derived from electronic trace data and how we pre-process those data to reflect social science constructs (Goggins et al., 2013). In later sections, we present the conceptual LoO used in our analysis of electronic trace data in each case. In doing this, we reveal patterns of abstraction across technologies and insights about the abstractions of organization in each individual social technology.

The base abstractions that each social technology has in common are participation and affiliation, which have specific, reflexive relationships in each social technology. At the most abstract level, influence may be understood as a function of reach (number of affiliations) and participation (social technology specific). The greater a person's relationships of affiliation, the more likely the person is to have influence, following from the basic theory of large numbers in statistics (Engländer, Harris, & Kyprianou, 2010), which specifies the aggregation of reach as "the rich get richer." The two abstractions are reflexive. The more a person is seen or shared, the more people he or she is exposed to. This communication-style influence reflects existing theories of communication (McLuhan & Fiore, 2005). Participation has a greater effect on future affiliations than affiliation has on participation when examining the LoE of influence. This follows logically from our understanding of the concepts above; the more people who follow a person's work, or are aware of it, the greater the opportunity for influence and the generation of additional followers (affiliations). Direct influence, however, is garnered strictly from participation. The affiliation relationships can increase the weight of influence, but without participation, there is no influence.

LoO for Influence LoE on Facebook

Our research team has examined more than one million interactions on Facebook, focused on political groups and individual users (Mascaro, Novak, & Goggins, 2012). Our goals were to understand how ideas propagate through Facebook and how civil discourse and political influence during a political campaign are operationalized during a U.S. election cycle. Figure 1 describes LoOs for interactions on Facebook that emerge from our study.

Participation and affiliation are base abstractions in Facebook, Twitter, and GitHub. Distinct interconnections between abstractions exist in each case. For Facebook, reciprocal affiliation connections between people necessary for information visibility form a closed loop; unless posting (participation) settings are changed to "public," I

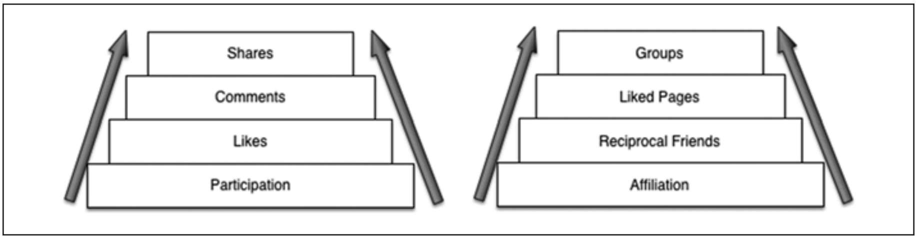

Figure 1. Levels of organization related to explaining influence on Facebook.
Note. The bottom level is viewed as the base abstraction. As we move up the pyramid, additional data points are additive to the construct of influence, creating a more context-specific picture of influence components.

cannot see your posts if we are not friends. The basic form of participation is the "like," which signals agreement. Here, influence is measured from the perspective of agreement. There is a low level of influence (we already agree) and, therefore, opinions are not changed so much as reinforced. In our examination of comments on Facebook, our prior work finds disagreement to be most pronounced on political topics.

In cases where issues are debated, there is often frequent exchange in a comment thread underneath a post, where people voice agreement and disagreement and exercise persuasion. These acts of discourse on social media are influence activities. Higher participation in threads connects to greater exercise of influence, however, not all comment threads have this characteristic—the context matters. If a person comments on a song that another person likes, there is less influence than there is simple, social engagement.

Influence in social technologies is also a function of how many people see a post. Researchers outside of Facebook's internal operation have little insight into these data, but we can discern that if a post or topic is shared from one group or user's page to others, the information is more visible. Consequently, sharing posts is an act that may spread the intended influence to a wider audience. As we move up the participation pyramid, the growing aggregation of influence acts on Facebook leads to a more complete measure of overall influence.

LoO for Influence LoE on Twitter

Twitter, in contrast with Facebook, enables asymmetric connections between participants. Users can follow other users without being followed. This asymmetry alters influence dynamics on Twitter in such a way that specific mechanisms for measuring influence are differently grounded in the LoO of Twitter data. Recall that these LoOs are specifically oriented to the construct of influence, which is the critical step we are taking. Figure 2 illustrates the LoOs for Twitter.

First, affiliation on Twitter follows two paths derived from the navigation of Twitter users. The first path is similar to Facebook, involving a follower relationship.

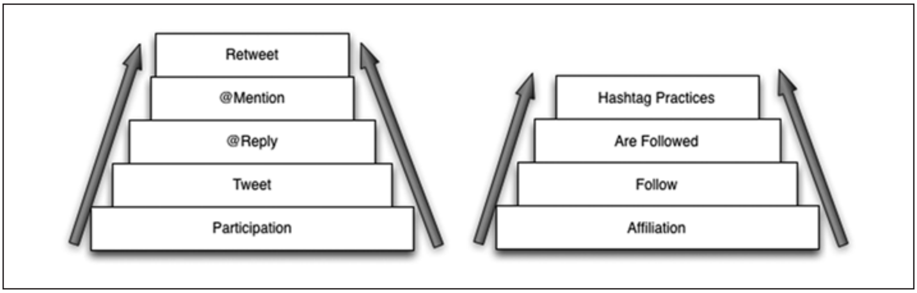

Figure 2. Levels of organization related to explaining influence on Twitter.
Note. The bottom level is viewed as the base abstraction. As we move up the pyramid, additional data points are additive to the construct of influence, creating a more context-specific picture of influence components.

Following a person gives him or her a degree of influence. The more followers an individual has, the larger the audience for the individual's work, creating a potential for greater influence. We follow political groups on Twitter and have analyzed congressional elections from the United States in 2012 (Black, Mascaro, Gallagher, & Goggins, 2012) to arrive at the influence abstraction in Figure 1.

The third path of affiliation on Twitter involves hashtag use. Tweets marked with a hashtag contribute to topically focused discussions, and others may follow discussion around the hashtag. During the 2012 U.S. election, key hashtags related to candidates included #tcot (top conservatives on Twitter) and #p2 (a corresponding liberal hashtag). This is the finest level of abstraction in our model because hashtag navigation is topical and orthogonal to social navigation.

Both hashtag and following affiliations allow the measurement of the reach of a particular tweet, the base unit of participation on Twitter. Participation is the key measure of actual influence; if few people follow you or use the hashtags you create, influence is low. Recently, "favoriting" of tweets has grown in popularity. Although our empirical work does not incorporate favoriting, we speculate that this behavior is a weaker component of the participation LoO than retweeting.

The LoO for influence reflects the components in Figure 2. The tweet is the core LoO. The @reply is a direct, public indicator from one user to another, signaling an explicit measure of influence. In this case, communication is direct, and the influence exerted is specific to that conversation, although others may be watching. In addition, you may also be mentioned in somebody else's tweet. If you are a political candidate, people who follow you see tweets where you are mentioned. The audience for these tweets is then potentially greater. These relations indicate potential influence.

Our analysis of the 2012 U.S. election suggests that the retweet signaling marker most strongly indicates influence. The retweet is an act where a follower shares an individual's tweet with their followers. The more retweets, the greater one's reach and the higher the likelihood of influencing a large number of people. How might we measure the influence of a cascade of tweets over time?

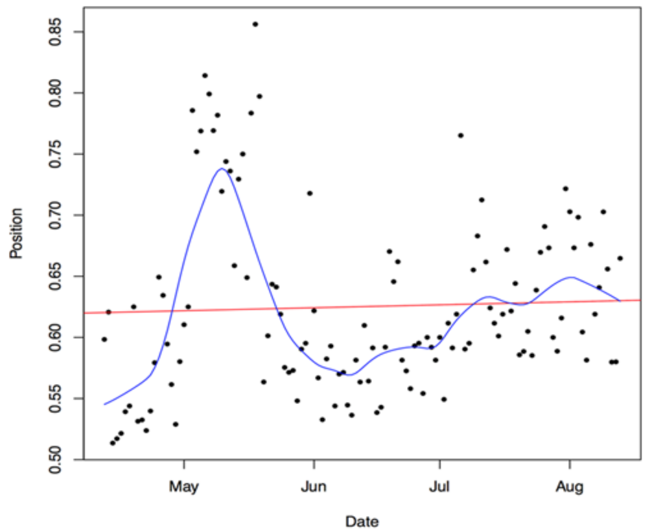

Figure 3. Hashtag #economy median position by day.
Note. The bottom level is viewed as the base abstraction. As we move up the pyramid, additional data points are additive to the construct of influence, creating a more context-specific picture of influence components.

One of our methods for discerning shifts in discourse due to retweets is to attend to hashtag use within a tweet corpora around that hashtag over time. When a tweet is marked with a hashtag (#economy), we view this as an act of classification, a less direct form of engagement than conversation. When hashtag use shifts toward the middle of the tweet on average, the likelihood that the construct is embedded in a conversation or debate is greater, based on our analysis of content. As an example, in Figure 3, we illustrate how the hashtag #economy shifted during the middle of the 2012 U.S. presidential election cycle. We suggest that shifts in a hashtag position are an indicator of influence on Twitter worth pursuing.

The blue line in Figure 3 is a LOESS curve, a statistic useful for describing shifts in multivalent data over time. The period in early May 2012, where tags shift toward the end of tweets in aggregate, signals that people use #economy more to classify tweets than to discuss the economy, and this time period corresponds with a more political, less sober discussion of economic issues in the United States.

LoO for Influence LoE on GitHub

Purposes for participation on GitHub are fundamentally different from those on Facebook or Twitter. Open source software projects are driven by the production of a working product more than idea exchange. The nature of influence focuses more on

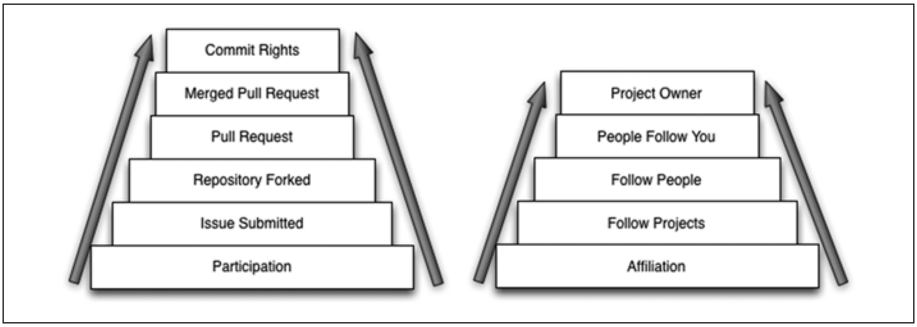

Figure 4. Levels of organization related to explaining influence on GitHub.

contribution to a codebase than on conversation around it. Contribution creates credibility, and high contributors may be granted specific roles on a project, reflecting appreciation for contributions and signaling influential individuals. Nonetheless, LoO abstractions of participation and affiliation are present that are central to understanding how participatory influence is operationalized on GitHub. Figure 4 illustrates the LoOs identified as related to influence on GitHub.

As with Twitter, one may follow a particular project or individual on GitHub. However, the nature of the information obtained is differently focused on alerts related to code committed by individuals or to a particular repository. Consequently, the nature of influence related to affiliation on GitHub is somewhat lower. GitHub emphasizes work, and the work is focused on participation. The exception to this is when one is an owner of a particular repository. This is a status indicator; ownership of an active repository indicates substantial influence on a project.

On GitHub, participation follows a trajectory from talk-based participation, through contributions directly to code, which we refer to as work-based participation. The LoOs on GitHub correspond with understood practices in open source software projects, where members move from the periphery into the role of core contributors in some cases. The more code one commits, the greater one's influence in a particular open source project (Crowston, Wei, Howison, & Wiggins, 2012).

The specific organization on GitHub, however, enables a wider degree of both participation and opportunity for influence, which differs slightly from traditional open source software projects. Anyone can copy (fork) a GitHub repository, change the code in their copy, and submit code changes via a "pull request," which is then debated among project participants. Influence is more directly measurable on GitHub when contrasted with Facebook and Twitter—if your code is committed, that is a signal that you have influenced the other participants to accept your contribution as "worthy" or "useful."

In our mixed methods analysis of 32 projects on GitHub, we have found connections between influence and the electronic trace data available through the GitHub API (Application Program Interface). Figure 4 shows that participation begins with the submission of an issue. If an issue results in code, there is greater influence than if it does not. As we cumulatively move up the LoO for GitHub participation in Figure 4, an individual's influence on a project becomes explicitly greater. Forking a repository is a work act; submitting a pull request engages discussion with other project members; and merging the pull request with code indicates that a participant is influencing the project. Individuals with commit privileges on a repository are strongly influential as they become a decision maker for submitted pull requests.

Toward a Context-Sensitive Theory of Influence in Social Technologies

Influencing each other is a core human activity, however, little is understood about influence in social technologies. To understand how theories of influence from the physical world interact with influence practices in social technologies, future research examining influence in social technologies can (and, we argue, should) be more grounded through the abstractions (LoO) that we describe here. Through this foregrounding of epistemological commitments, researchers can advance a set of theories related to influence in social technologies to motivate targeted empirical work.

The LoO abstractions for influence described here are not explicitly derived from constructs in these social technologies. The LoOs we present are synthesized from empirical analysis of influence in Facebook, Twitter, and GitHub and reflect the explicit epistemological commitments noted above.

Our operationalization of influence through the philosophical lens of levelism reflects an attempt to connect the philosophy of science to the analysis of electronic trace data. This enables researchers to be explicit about how these commitments foreground their epistemological commitments in the analysis of trace data from social technologies. An alternate and more common strategy would be to take the electronic trace data at face value. This would lead to operationalizing connections within trace data without being explicit about which connections matter for a particular, desired explanation. This desired explanation would then be built on opaque ontological commitment, making comparisons across social technologies and sociotechnical contexts difficult.

In our empirical work, we connect our analysis of influence through explicit abstractions of influence (LoEs) derived from interviews and content analysis. This frames analysis of electronic trace data. We argue that these kinds of commitments are both necessary and rare in the present-day analysis of influence in social technologies. Our foundational, empirical work and the organizing abstractions described here make the connections found in trace data explicit, which supports development of theories of influence vis-à-vis testing of these abstractions in future studies.

The Importance of Looking Across Social Technologies

There is a great deal of empirical work around emerging social technologies. Additional development of theories of influence that account for multiple, specific social technologies is needed to create theories of influence specific to the universe of social technologies. These limitations can be attributed to the substantial technical and social scientific work required to analyze just one social technology. Our publications and reflections across several social technologies enable our thinking at a more abstract level. Future limitations will connect to growing social awareness of the privacy limitations of social technologies and government surveillance of online activities. The role of voyeurism and discussion about how watching and being watched is connected to influence across technologies should be a growing area of theory and policy development.

We introduce the core abstractions of participation and influence as common levels of organization in each social technology and then operationalize these abstractions using lower level constructs derived from each social technology. From this work, we seek a foundation on which to build theory.

Our empirical work has led us to develop ontologies and methodological approaches to understand how to ensure coherence between network analytic constructs of degree and betweenness centrality and the electronic trace data that these systems produce. From this work, we derive and present abstractions as outlined here. We think that there is a need for the reflexive development of theories of influence that incorporate the LoOs here and more traditional theories of organization and influence.

We argue that there are common abstractions for organizing electronic trace data from social technologies and provide examples from our empirical work. There are also distinct characteristics for each social technology that render them different in purpose and, for researchers, likely suggest differences in how we frame influence. Facebook connects us to those with whom we communicate through affiliation relations at all times; it is core to the design. Twitter, in contrast, emerges more as a form of "participatory mass media," with multiple tracks for affiliation and the capacity for large-scale sharing of tweets. GitHub is a third genre of social technology, one focused on the completion of work, where influence is indicated more by what is done than what is written. Although the abstractions are common at some level, the nature of participation in each of these technologies is decidedly different.

Connecting to Theories of Influence From the Physical World

Present-day influence theories exist in small group literature, organizational literature, and communications literature. Each theoretical perspective should bring a systematic methodological approach for making sense of trace data, built on a structuring of trace data that is clearly tied to the phenomena that researchers explicate.

We argue that discrete social constructions of the small group and the organization are less clearly defined in social technologies than they are in the physical world. This compels research on influence to look across the levels that we are trained to examine. The LoO abstractions presented here make that decision explicit for future empirical research. Small group literature focuses on the various reasons that groups come together in the physical world, especially for social and task-focused activities (McGrath & Hollingshead, 1993). Organizational theories, particularly structuration theory, address concerns about the reflexive nature of human and organizational engagement through technology (Orlikowski & Barley, 2001). Future research might consider synthesizing theory across both levels of influence in social technologies and organizational levels.

Declaration of Conflicting Interests

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

Funding

The author(s) received the following financial support for the research, authorship, and/or publication of this article: Parts of the research were funded by the National Science Foundation, VOSS award, "Toward a Context Adaptive Theory of the Relationship Between Structural Fluidity and Virtual Organization Performance." The authors also acknowledge past and present members of the Group Informatics Lab at Drexel Univerity and the University of Missouri, for their contributions to discussions leading to this article. Also, the Consortium for the Science of Sociotechnical Systems summer institutes substantially influenced this paper's direction.

References

Bansler, J. P., & Havn, E. (2006). Sensemaking in technology-use mediation: Adapting groupware technology in organizations. *Computer Supported Cooperative Work*, *15*(1), 55-91.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *The Journal of Political Economy*, *100*(5), 992-1026.

Bird, C., Rigby, P. C., Barr, E. T., Hamilton, D. J., German, D. M., & Devanbu, P. (2009). The promises and perils of mining Git. In *6th IEEE international working conference on mining software repositories* (pp. 1-10). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5069475

Black, A., Mascaro, C. M., Gallagher, M., & Goggins, S. P. (2012). Twitter zombie: Architecture for capturing, socially transforming and analyzing the Twittersphere. In *Proceedings of the 2012 ACM group conference* (pp. 229-238). *Sanibel Island, FL: ACM*.

Boyd, D. (2010, April 29). *Privacy and publicity in the context of Big Data*. Paper presented at the WWW conference, Raleigh, NC.

Chang, H. C. (2010). A new perspective on Twitter hashtag use: Diffusion of innovation theory. *Proceedings of the American Society for Information Science and Technology*, *47*(1), 1-4.

Crowston, K., Wei, K., Howison, J., & Wiggins, A. (2012). Free/Libre Open Source Software development: What we know and what we do not know. *ACM Computing Surveys, 44*. doi:10.1145/2089125.2089127

Dabbish, L., Stuart, C., Tsay, J., & Herbsleb, J. (2012). Social coding in GitHub: Transparency and collaboration in an open software repository. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (pp. 1277-1286). *New York, NY: ACM*. doi:10.1145/2145204.2145396

De Souza, C. S., & Preece, J. (2004). A framework for analyzing and understanding online communities. *Interacting With Computers*, *16*(3), 579-610.

Dourish, P. (2004). What we talk about when we talk about context. *Personal and Ubiquitous Computing*, *8*, 19-30.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, *14*(4), 532-550.

Engländer, J., Harris, S. C., & Kyprianou, A. E. (2010). Strong law of large numbers for branching diffusions. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, *46*, 279-298.

Erickson, T., & Kellogg, W. A. (2000). Social translucence: An approach to designing systems that support social processes. *ACM Transactions on Computer-Human Interaction*, *7*, 59-83.

Finley, K. (2011). *Github has surpassed Sourceforge and Google Code in popularity*. Retrieved from http://readwrite.com/2011/06/02/github-has-passed-sourceforge#awesm=~oxkoadKkSk6gAo

Floridi, L. (2008). The method of levels of abstraction. *Minds and Machines*, *18*(3), 303-329.

Gil de Zúñiga, H., Jung, N., & Valenzuela, S. (2012). Social media use for news and individuals' social capital, civic engagement and political participation. *Journal of Computer-Mediated Communication*, *17*(3), 319-336.

Gilbert, E., & Karahalios, K. (2009). Predicting tie strength with social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 211-220). Retrieved from http://dl.acm.org/citation.cfm?id=1518736

Goggins, S. P., & Mascaro, C. M. (2011). *Social media discourse and culture: A proposal for comparative informatics research*. Paper presented at the 2nd annual workshop on comparative informatics, Copenhagen, Denmark.

Goggins, S. P., Mascaro, C. M., & Valetto, G. (2013). Group informatics: A methodological approach and ontology for understanding socio-technical groups. *JASIS&T*, *64*(3), 516-539.

Golder, S. A., & Yardi, S. (2010). Structural predictors of tie formation in Twitter: Transitivity and mutuality. In *Proceedings of the 2010 IEEE Second International Conference on Social Computing* (pp. 88-95). Washington, DC: IEEE Computer Society. doi:10.1109/SocialCom.2010.22

Grabowicz, P. A., Ramasco, J. J., & Eguiluz, V. M. (2012). *Dynamics in online social networks*. Retrieved from http://arxiv.org/abs/*1210*.0808

Howison, J., Wiggins, A., & Crowston, K. (2012). Validity issues in the use of social network analysis with digital trace data. *Journal of the Association of Information Systems*, *12*(2). Retrieved from http://aisel.aisnet.org/jais/vol12/iss12/2/

King, G. (2011). Ensuring the data rich future of the social sciences. *Science*, *331*(6018), 719-721.

Lampe, C., Ellison, N., & Steinfield, C. (2006). *A face(book) in the crowd: Social searching vs. social browsing*. New York, NY: ACM.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., …Van Alstyne, M. (2009). Life in the network: the coming age of computational social science. *Science*, *5915*(323), 721-723. doi:10.1017/S1049096510001873

Malhotra, Y., & Galletta, D. (2003). A multidimensional commitment model of volitional systems adoption and usage behavior. *Journal of Management Information Systems*, *22*(1), 117-151.

Mascaro, C. M., Novak, A. N., & Goggins, S. P. (2012). The daily brew: The structural evolution of the coffee party on Facebook during the 2010 United States midterm election season. *Journal of Information Technology and Politics*, *9*(3), 234-253.

McGrath, J. E., & Hollingshead, A. B. (1993). *Groups interacting with technology: Ideas, evidence, issues and an agenda*. Thousand Oaks, CA: SAGE.

McLuhan, M., & Fiore, Q. (2005). *The medium is the massage*. Berkeley, CA: Gingko Press.

Miller, G. (2011). Social scientists wade into the tweet stream. *Science*, *333*(6051), 1814-1815.

Orlikowski, W. J., & Barley, S. R. (2001). Technology and institutions: What can research on information technology and research on organizations learn from each other? *MIS Quarterly*, *25*(2), 145-165.

Steinfield, C., Ellison, N. B., Lampe, C., & Vitak, J. (2012). Online social network sites and the concept of social capital. *Frontiers in New Media Re-search*, *15*, 115.

Urry, J. (2000). *Sociology beyond societies: Mobilities for the twenty-first century*. New York, NY: Routledge.

Valenzuela, S., Park, N., & Kee, K. F. (2009). Is there social capital in a social network site? Facebook use and college students' life satisfaction, trust, and participation. *Journal of Computer-Mediated Communication*, *14*(4), 875-901.

Watts, D. (2007). A twenty-first century science. *Nature*, *445*(7127), 489.

Wellman, B., Haase, A. Q., Witte, J., & Hampton, K. (2001). Does the Internet increase, decrease, or supplement social capital? Social networks, participation, and community commitment. *American Behavioral Scientist*, *45*(3), 436-455.

Yardi, S., & Boyd, D. (2010). Dynamic debates: An analysis of group polarization over time on Twitter. *Bulletin of Science, Technology & Society*, *30*(5), 316-327. doi:10.1177/0270467610380011

Yuan, C., & Gay, G. (2006). Homophily of network ties and bonding and bridging social capital in computer-mediated distributed teams. *Journal of Computer-Mediated Communication*, *11*, 1062-1084.

Author Biographies

Sean Goggins is an Assistant Professor at Missouri's iSchool and the University of Missouri Informatics Institute. He teaches, publishes, and conducts research on the uptake and use of information and communication technologies by small groups in medium to large scale socio-technical systems, from Facebook, to online course systems. Sean conceptualizes "group informatics" as a methodological approach and ontology (Goggins et al., 2013) for making sense of the interactions between people in medium to large scale social computing environments. His research examines the information behavior, knowledge construction, identity development, performance and structural evolution of small, online groups. The long tail of social computing bounds this work; large scale communities are approached from the bottom up in group informatics research. By understanding small groups of online learners, librarians or software engineers (to name just a few), and analyzing traces of their interactions, Group Informatics brings semantics to the clusters discovered in large scale social computing endeavors.

Eva Petakovic is a research assistant at the University of Missouri Informatics Institute who works with Dr. Sean P. Goggins on social media and technology research. Her research interests are forming in the area of technologically mediated societies.