Journal of the Association for Information Systems
JAIS

Research Practice

Validity Issues in the Use of Social Network Analysis with Digital Trace Data

James Howison
University of Texas at Austin
jhowison@ischool.utexas.edu

Andrea Wiggins
Syracuse University
awiggins@syr.edu

Kevin Crowston
Syracuse University
crowston@syr.edu

Abstract

There is an exciting natural match between social network analysis methods and the growth of data sources produced by social interactions via information technologies, from online communities to corporate information systems. Information Systems researchers have not been slow to embrace this combination of method and data. Such systems increasingly provide "digital trace data" that provide new research opportunities. Yet digital trace data are substantively different from the survey and interview data for which network analysis measures and interpretations were originally developed. This paper examines 10 validity issues associated with the combination of digital trace data and social network analysis methods, with examples from the IS literature, to provide recommendations for improving the validity of future research.

***Keywords**: Social Network Analysis, Validity, Methods, Digital Trace Data.*

* Youngjin Yoo was the accepting senior editor. This article was submitted on 15th October 2008 and went through four revisions.

Volume 12, Issue 12, pp. 767-797, December 2011

Validity Issues in the Use of Social Network Analysis with Digital Trace Data

1. Introduction

There is an exciting natural match between Social Network Analysis (SNA) and the growing phenomenon of social interaction through digital platforms and technologies, from online communities to corporate information systems (Agarwal, Gupta, & Kraut, 2008). This match offers a combination of exciting phenomena, interesting research questions, appropriate analysis techniques, and the availability of copious data. Agarwal et al. (2008) put it thus: "Most transactions and conversations in these online groups leave a digital trace…this research data makes visible social processes that are much more difficult to study in conventional organizational settings." The availability of such trace data, together with exciting domains and an appropriate analysis technique, form a golden opportunity for research, perhaps even a "21st Century Science" (Watts, 2007).

The discipline of Information Systems has not been slow to recognize and explore this natural match. Rice (1990) laid out an early case explicitly:

> *The fact that CMC systems can unobtrusively collect data on usage, flows, and content from a full census of users provides researchers with new opportunities for understanding the application, management, and consequences of such systems. A theoretically appropriate analytical approach is network analysis of CMC system data (p. 643).*

Information Systems researchers have embraced this opportunity, undertaking innovative research on a variety of topics, including group cohesion (e.g., Hahn, Moon, & Zhang, 2008), trust (e.g., Ridings, Gefen, & Arinze, 2002), knowledge generation (e.g., Wasko & Faraj, 2005), information diffusion (e.g., Hinz & Spann, 2008), and productivity (e.g., Aral, Brynjolfsson, & van Alstyne, 2006) in a wide range of domains, including virtual collaborations (e.g., Ahuja & Carley, 1999), Wikipedia (e.g., Kane, 2009), free/libre open source software development teams (e.g., Wu & Tang, 2007), electronic commerce (e.g., Bampo, Ewing, Mather, Stewart, & Wallace, 2008), and corporate workflow (e.g., Brynjolfsson, Malone, Gurbaxani, & Kambil, 1994; Robey, Vaverek, & Saunders, 1989).

Researchers in cognate disciplines are similarly recognizing the potential of this match, as Kleinburg (2008, pp. 66–67) writes:

> *Collecting social-network data has traditionally been hard work, requiring extensive contact with the group of people being studied; and, given the practical considerations, research efforts have generally been limited to groups of tens to hundreds of individuals. Social interaction in online settings, on the other hand, leaves extensive digital traces by its very nature…we can replay and watch…the ephemeral dynamics of ordinary life, now made visible through their online manifestations. As such, we are witnessing a revolution in the measurement of collective human behavior.*

A measurement revolution is an exciting time, but it is also a time that calls for reflection; with opportunities come risks, especially when methods developed in one context are applied in new contexts. In particular, the underlying assumptions of traditional social network analysis methods have not often been examined in detail when using digital trace data. Indeed, a review of reliability and validity of measures of information structures addresses this type of data only briefly and uncritically (Zwijze-Koning & de Jong, 2005). This situation is reason for concern, as the available data and the kinds of structures they represent differ in key respects from the data and structures addressed in earlier social network studies. Failure to address these differences can threaten the validity of network measures, and can undermine the whole "chain of reasoning" (Hume, 2000, sec. Advertisement) that leads to reported results using SNA with digital trace data. If this exciting combination of phenomena, research questions, data, and method is to reach its promise, these issues must be addressed.

This paper presents a series of decisions researchers have to make in executing a network study using digital trace data. For each decision, we highlight threats to validity, placing them in the context of existing validity frameworks commonly used in IS. We discuss the source of these threats and provide illustrations of potential mistakes drawn from existing IS literature. We also showcase studies that have dealt well with the threats. Finally, for each issue we provide a set of recommendations for how to address the issue in research and review.

1.1. Defining Social Network Analysis

SNA is not a theory *per se*; it is a set of analysis techniques (thus, SNA rather than SNT). Various substantive theories (e.g., Monge & Contractor, 2003) focus attention on networks in different settings, motivating the use of graph network analysis techniques, but these theories and the analysis techniques are conceptually distinct. There is a growing body of work that countenances building general network theory, often called “network science,” (e.g., Committee on Network Science for Future Army Applications, National Research Council, 2005; Kilduff & Tsai, 2003) but this project is not complete and, in any case, the techniques of SNA are frequently used outside such theoretic perspectives.

As a result, it is at best incomplete to speak of SNA findings, just as it would be to speak of regression findings. Indeed, the use of SNA techniques parallels those of other such quantitative techniques. For analysis, a set of relationships is represented as a mathematical structure (a graph) composed of nodes and links, often encoded as an interaction matrix. Thus, the use of SNA requires the network to have been measured as a graph, just as the use of conventional statistical techniques requires that constructs of interest be measured as series of variables. Given a graph or interaction matrix, calculations can be made of individual-level scores for the structural position of nodes, such as various individual scores for network centrality, as well as measures providing overall summaries of structural characteristics for the whole network, such as network density or centralization. The application of these techniques is conceptually similar to the statistical computation of an individual score, such as a z-score, to show an individual's relative position in a distribution, or a summary statistic, such as a mean or standard deviation, to summarize an entire sample.

Just as statistical analysis techniques like averaging and finding standard deviations can be applied to data representing a wide diversity of constructs, SNA techniques can be applied to networks built from data representing diverse kinds of nodes and links, each with different theoretical characteristics. Those characteristics bear directly on the validity of interpretations. The goal of the paper is to consider how novel kinds of data raise different questions to be addressed by researchers.

1.2. Defining Digital Trace Data

This paper considers validity issues in network analysis when working with digital trace data. We define digital trace data as records of activity (trace data) undertaken through an online information system (thus, digital). A trace is a mark left as a sign of passage; it is recorded evidence that something has occurred in the past. For trace data, the system acts as a data collection tool, providing both advantages and limitations. The task for using this evidence in network analysis is to turn these recorded traces of activity into measures of theoretically interesting constructs.

All trace data, not just digital trace data, has three characteristics that underlie many of the issues discussed in this paper: 1) it is found data (rather than produced for research), 2) it is event-based data (rather than summary data), and 3) as events occur over a period of time, it is longitudinal data. In each aspect, such data contrasts with data traditionally collected through social network surveys and interviews.

First, trace data are **found** data in the sense that they are a by-product of activities rather than produced by a designed research instrument. Wikipedia was not designed to test theories about knowledge production, nor are corporate email systems designed to collect research data. This origin contrasts with social network surveys or interviews that are specifically designed to produce data for research. Trace data, as found data, must be adapted for research purposes. Indeed such data might even prove to be more useful for some research questions for that very reason, once the validity concerns discussed in this paper are addressed.

Second, trace data are **event-based data**, rather than summary-based data. In a traditional SNA survey, researchers typically ask directly about social relationships, relying on the respondents to recall and interpret their own interactions to summarize a social relationship. By contrast, with trace data, researchers themselves must make the move from evidence to measure and from event to relationship.

Of course, some events (and records of events) provide better evidence of a social relationship than others. At one end of this spectrum, some events, by their mere occurrence, provide summarized evidence of a social relationship. A wedding is an event, but is itself an expression, even an enactment, of a social relationship and is, therefore, strong evidence for a past and future social relationship. In a similar way, the act of “friending” someone in an online social network is both an event leaving a trace and a signification of some type of social relationship. However, what can be inferred from an event depends on the meaning the participants and their social context give it. Nonetheless, in some circumstances, by undertaking the action leading to the record, the participants are explicitly attempting to signify some social relationship.

Much trace data, however, does not have such a signifying quality: a reply to an email on a mailing list seems unlikely to be an attempt to summarize a social relationship. Yet, as a trace of activity and a type of interaction, it may provide evidence about a social relationship; careful research may make inferences without relying on the actors’ direct understanding of their social relationships. Many of the issues in this paper stem from this understanding of the task facing researchers: Trace data show evidence of the “raw material” of social relationships, so the research task is to understand what can be inferred about higher-order constructs from the existence of the trace data.

The final key characteristic of trace data are that they are **longitudinal data**, because the events that make it up occur over time. To apply network analysis techniques, the multiple events have to be aggregated to produce evidence of a network structure. Surveys typically ask respondents to report on a period of time, up until the point of the survey; but with trace data, researchers have to make decisions about how to deal with converting events that occur over time into networks.

In defining trace data, it is worth noting the relationship between trace data and archival data. Archival data are those that are stored in and retrieved from an archive, rather than collected anew. Such archives could contain both trace data and data that represents participants’ summaries of their social relationships (i.e., not trace data). For this reason, one can say that all trace data are archival, but not all archival data are trace data. By using the term trace data, we seek to emphasize that what is left in the archives is distinct; it is a trace of activity, indirect evidence for, rather than a direct measure of, a social relationship. Patent citations are a good example: The existence of a citation is direct evidence of a citing event, an author choosing to insert a citation into a patent. Converting from knowledge of this event into a construct such as knowledge flow may be a reasonable interpretation of the evidence, given an appropriate theory, but it is an interpretation nonetheless, and it ought to be argued as valid.

The second part of the definition of **digital trace data** is that the data are both produced through and stored by an information system. Not all trace data are digital in this sense, including patent citations. Moreover, trace data could be produced through direct observation. An example might be watching people in a lunchroom or constantly recording audio feeds that are then processed to produce network maps. In Information Systems research, however, the growth of online interaction has lead to a marked increase in the availability and research use of explicitly *digital* trace data. In this respect, the involvement of a specific communication or information system is important. As we consider the issues below, we highlight those that are likely present with all trace data and those that stem more specifically from the involvement of an information system.

Trace data are not new in SNA, but until recently data from questionnaires and interviews have been strongly preferred, and trace data relied upon only when these have not been possible (Wasserman & Faust, 1994). This preference is reflected in the articles in the key SNA journal, *Social Networks*. Our examination shows that there are almost no articles that make use of trace data alone (with Adamic and Adar (2005) a recent exception; they rely only on digital trace data).

The far more widely used survey methods, such as name generators and social network interviews, have developed their own literature of validity. Marsden (1990), for example, shows that people are notoriously poor at reporting discrete interactions but generally good at recalling long-term social structures. Other researchers have considered the differences between perceived networks and actual behavior (e.g., Kilduff, Crossland, Tsai, & Krackhardt, 2008), describing the limits of working with survey data to predict actual behavior. This paper is a step toward developing a corresponding understanding of the validity issues posed when working with trace data, especially in its digital form.

1.3. Defining Validity

Validity is a concern in all research; it concerns the approximate truth of an inference. As Sechrest (2005) notes, "Validity must be considered to inhere in a system or process of which the instrument itself is only a feature." The relevant system in this context is the researcher's theoretical context, which first suggests theoretical constructs to be measured. To argue that the measurement is valid, the researcher builds a chain of reasoning linking construct to data. This chain must run logically in both directions, from data to construct and construct to data.

The Information Systems field has found the validity frameworks developed by Cook and Campbell (1979) and Shadish, Cook and Campbell (2001) particularly useful for understanding validity. These frameworks divide validity issues into four categories spanning the chain of reasoning in research: construct validity, statistical conclusion validity, internal validity, and external validity. Construct validity refers to the extent to which operationalizations (or measures) validly approximate theoretical constructs. Statistical conclusion validity refers to the extent to which statistics validly support the inference that measures co-vary. Internal validity reflects the extent to which the inference that such covariance is due to causality is valid. External validity refers to the validity of inferences about the extent to which such cause-effect relationships hold in different research settings (often referred to as generalizability).

The analysis of validity is not a formulaic exercise. Indeed, the Cook and Campbell (2001) validity framework is, in the words of its authors, "practical only" and the categories are derived from "their apparent correspondence to four major decision questions that the practicing researcher faces." (Shadish et al., 2001, p. 39). These categories align most clearly with experiment-based research designs, though they have been extended to cover quasi-experimental approaches as well. However, research using SNA with digital trace data employs a wide variety of approaches, only some of which naturally resemble experimental structures. Therefore, in the spirit of Cook and Campbell (2001), we frame our study of validity issues with respect to the decisions practicing researchers must make, relating to the Cook and Campbell validity framework as appropriate. The issues raised below relate to Cook and Campbell's categories of construct, internal, and statistical conclusion validity. We do not deal explicitly with issues of external validity, since we do not find that working with digital trace data raises particular external validity issues beyond those relevant and important to research in general.

2. Alignment along the Chain of Reasoning

To ensure the validity of network research, researchers must think carefully about the network process at play in their theory, consider appropriate network measures, identify appropriate operationalizations of nodes and ties in the context of their data, and so, connect to measures and constructs, iterating through the chain of reasoning until it is cohesive, as shown in Figure 1. At the top of Figure 1 is a summary of the abstract chain of reasoning; at the bottom are two examples. Each link in this chain has validity implications, and it is around these links that we organize the remainder of this paper.

In practice, the process of achieving alignment between a theoretical context and the chain of reasoning underlying valid measurement is an iterative one, most likely involving multiple adjustments and decisions and revisiting these to achieve a cohesive logic. Within the limits of this paper, however, we must present the issues in a linear fashion. We do so according to a progression of reasoning from data to construct, though we do not suggest that research ought to be driven solely in this direction.

We start by considering an information system that creates digital trace data, raising issues of 1) system and social practice and 2) reliability. The next link we consider concerns transforming digital trace data into nodes & links, raising questions of 3) link types, 4) link intensity, and 5) missing links. Turning nodes and links into a network raises issues of 6) temporal aggregation; using that network to obtain a measure raises issues of 7) network tool effects and 8) temporal mismatch. Finally, aligning a measure and a construct raises 9) questions of data completeness and inference and 10) inappropriate importation of network measure interpretation. Of course, all of these decisions must be made in the context of some overall theory; therefore, we return to accomplishing theoretical cohesion across the full chain of reasoning in the Discussion section.

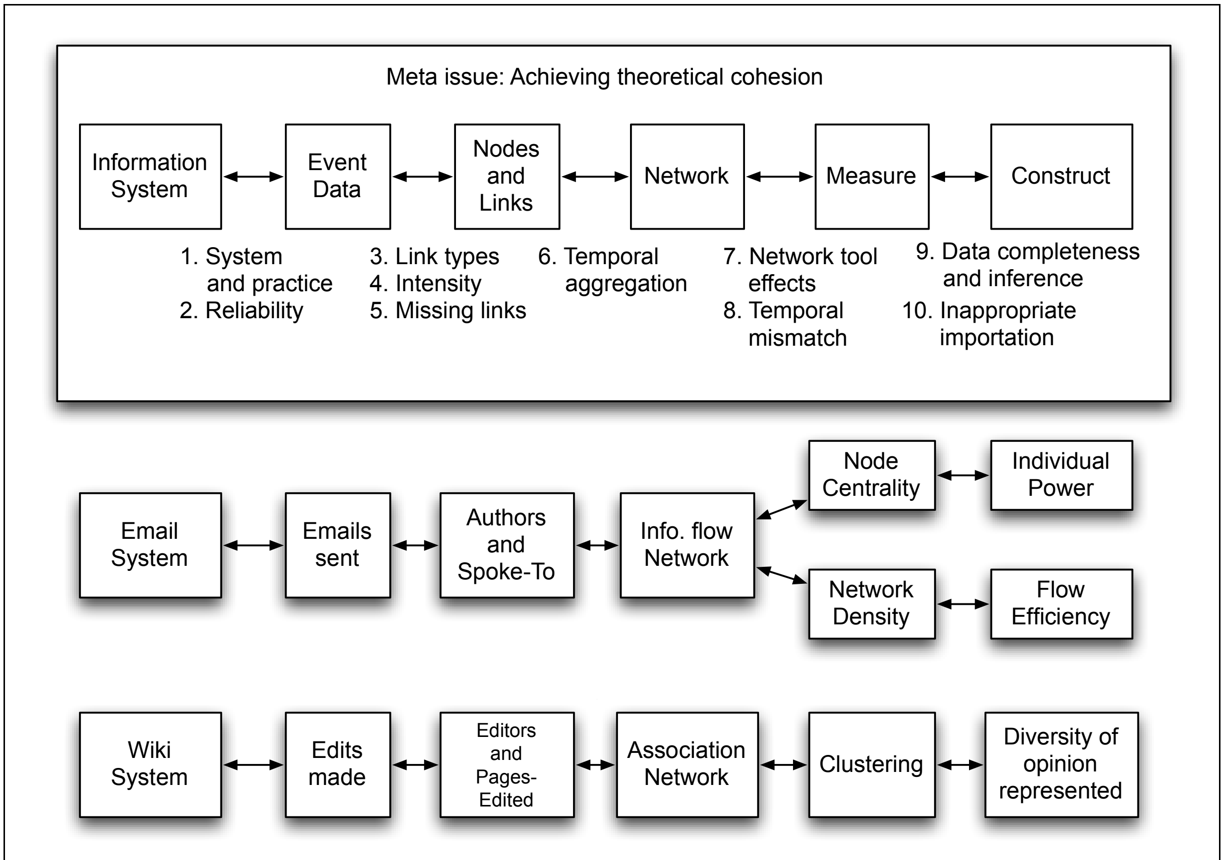

Figure 1. Links in the Chain of Reasoning and Validity Issues in Network Analysis with Digital Trace Data

2.1. Aligning Information System and Digital Trace Data

Information systems support an amazing variety of human activity, from work processes to social support, and are involved in collective activities that span a range of virtuality, from entirely online to those where the system is completely peripheral. It is surprising, therefore, that the specifics of the information system under consideration often do not appear in studies using digital trace data, as Orlikowski and Iacono (2001) note more generally. Moreover, it is a key understanding of Information Systems as a discipline that technologies are rarely used only as designed; design and use co-develop in a structurational process (Poole & DeSanctis, 1990) in which both the use of a technology and the technology itself change over time. This consideration gives rise to two key issues in using digital trace data for research: 1) understanding how the system is used in practice and how the specifics of the system impact behavior, and 2) how the system records behavior, especially over time, raising issues of data reliability.

Issue 1: System and Practice Issues

Databases of digital trace data typically come with system labels, such as “reply-to,” “friend,” “assigned-to,” and “member-of.” These evoke concepts of great interest to researchers. Yet the actual use, and therefore meaning, of these fields and records can be quite different from those concepts. For example, IBM’s JAZZ work collaboration system requires “membership” of a work team simply to view that team’s records; therefore, teams often have “members” who have done no work, in contrast to most conceptualizations of the role of a team member. In many community-based open source projects, to avoid discouraging others from working on a problem, the “assigned-to” field in a bug report is only filled out when a developer has finished the task (Howison, 2009), in contrast to the usual notion of proactive task assignment within work teams. Certainly these fields have some meaning, but it is problematic to assume an interpretation without an understanding of how the information system is used in practice. Since the information system, when interpreted, is also the measurement device for trace data, such misunderstandings can threaten construct validity, rendering data and measures derived from the data, at best, a poor proxy for the behavior and constructs of interest.

Moreover, the meaning of system-based interactions can change over time, even without obvious changes in the system or labels on the data. Long-term data are very useful, of course, but only if the researchers have adequately grappled with how they might have changed over time. For example, when using a data set based on software code change logs over 20 years (e.g., Merlo, Slaughter, & Francalanci, 2009), researchers should question whether it is reasonable to expect that the code version management tool has been used consistently (in ways that matter to the research) within the organizational context for two decades.

Similarly, it is important to understand how the use of the system is intertwined with unrecorded but relevant activity. Does the system capture nearly all of the interaction of the group, or does the group only use the system for a certain kind of interaction, or do they only use the system at particular times? What other systems are in use? Only with such understandings can the researcher grapple with the implications for their research context. It may, in fact, be of great interest to study and compare a “digital” network with a “face-to-face” network, but it would be a mistake to always reason on the basis that the digital network was the only source of interactions, as we discuss in detail in Issue 9, below.

System use waxes and wanes over time, especially as systems age and others come online. Researchers may need to understand such patterns to ensure that they have collected adequate data. For example, Wiggins, Howison, and Crowston (2008), in analyzing interactions on an open source bug tracking system, report one project in which hundreds of bugs had apparently been resolved within a few minutes. Detailed qualitative examination of this case revealed that the project had transferred bug reports from an old system to the one being analyzed via a bulk import. The transferred bug reports were, thus, stored with nearly identical open and close times. Including the data from this project in the analysis could have led to an incorrect inference regarding the causation of this burst of bug-fixing. If behavior is being measured over long periods of time, such changes in use can cause issues of construct validity through measurement error. If behavior is being measured in multiple short snapshots, such changes in use can cause issues of internal validity, since they may cause a false appearance of change in behaviors of interest (see Issue 7, below).

Understanding these issues, and the extent to which they matter for particular research questions, requires direct attention from researchers. We summarize the issues in Table 1 below (we will present a similar table for each subsequent issue). Clearly it is of great advantage to work directly with participants—through interviews, observation, and direct participation—to build a qualitative understanding of system use and how it fits into the overall interactions of a group. Geiger and Ribes (2011) call the process of taking digital traces and learning their meaning “inversion.” The event traces themselves are a particularly valuable point for developing understanding, since “documentary traces are the primary mechanism in which users themselves know their distributed communities and act within them.” (Geiger & Ribes, 2011, p. 1). For this reason, simply reading event records in sequence and working to reconstruct narratives can aid researchers significantly in understanding system use and establishing face validity in publications. Furthermore, the records themselves provide excellent anchors for interviews, helping participants recall specifics rather than generalities of their activity. Not all system-based research requires a full "trace ethnography" as called for by Geiger and Ribes, but studies using digital trace data as evidence ought to demonstrate to readers and reviewers that they have adequately grappled with issues of system use and its change over time.

Table 1. System and Practice Issues	
Decision	Do users, in fact, use the information system as measurement (often implicitly) assumes they do? How has that use changed over time?
Validity issue/type	Misunderstanding system use can lead to invalid interpretations of the data it collects. (Construct validity, measurement validity, statistical conclusion validity)
Cause	Systems are used in surprising and unexpected ways; database labels can take on different meanings in different contexts as well as change over time.
Examples	Wiggins et al. (2008)
Recommendations	• Gain intimate knowledge of the system, through interviews and participation, supported by the records themselves. Consider undertaking "trace ethnography" (Geiger & Ribes, 2011). • Demonstrate this familiarity with the system use context in publications, such as through illustrative narratives.

Issue 2: Reliability Issues from System Generated Data

On the surface, relying on a system to automatically collect data, as with digital trace data, would seem to ensure its reliability. Indeed Garton, Haythornthwaite, and Wellman (1997) go so far as to say "gathering data electronically replaces issues of accuracy and reliability with issues of data management, interpretation, and privacy." However, even if it can be established that the systems have been used in an adequately understood manner, to ensure reliability of measurements of digital trace data, it is essential to understand the processes by which the archives, and thus, data, are recorded and whether and how the system's recording processes have changed over time.

Unfortunately, a detailed examination of CMC systems may reveal numerous potential threats to reliability, such as inconsistent time zone management, server outages, and incomplete or inconsistent event logging, to name a few. For example, in a system that records email messages, times on the messages may be local time for the sender, local time for the server, GMT, or (in the worst case) some undecipherable combination. Resolving the question of what time a message was sent is difficult but necessary to reliably determine the order of messages or to aggregate the messages over time. More simply, a server crash may result in the loss of some data, likely with no explicit indications of a break in data integrity. A common problem that affects network research more specifically is that systems can have multiple system representations of a single user. Analyzing data that include these multiple representations results in splitting or merging network nodes in ways that might alter the whole network structure. Research on this topic has shown that the actual impact can be problematic and significant, but it depends on both the intended measure and the specific network topology (Franz, 1998), making general statistical control difficult.

Similar issues exist even with data that researchers do not collect themselves, such as database dumps provided by community systems. For example, the data provided to the Notre Dame Sourceforge Research Data Archive provide a convenient source of data about Sourceforge-based open source development projects (Gao, Antwerp, Christley, & Madey, 2007). Similarly, the Wikimedia Foundation has made available dumps of the database driving the Wikipedia system. Such data dumps can be used to build association networks based on membership or co-editorship, or communication networks drawing on issue trackers, forums, or talk pages (e.g., Kane, 2009).

However, the data in these systems exist to support the operation of the community, rather than being crafted for research. Therefore, pragmatic issues in operating the system will affect the reliability of measures constructed from this data, and often do so silently. For example, tables in many system databases are periodically purged to maintain a manageable size for running a website. This process results in database dumps with apparently extensive history that are actually truncated at an arbitrary date with no explicit record of such truncations. This problem is a very real issue in the (otherwise excellent) SRDA data set (Gao et al., 2007) where early dumps contain records that do not appear in later dumps, despite those later dumps including apparently full history tables. It is important to remember that the purpose of the Sourceforge database is running Sourceforge, not maintaining a full history of activity for researchers.

The English-language Wikipedia, as another example, has experienced issues with archiving due to its size, preventing full-text dumps from being made available for almost two years. The earlier history may be available from earlier dumps, but merging disparate, partially overlapping sources is quite difficult, particularly as incremental changes made over time may result in incompatible database schemas.[1] Similarly, systems that make usage-reporting data available may change their data sources or methods of calculation without notice, and almost undoubtedly without recalculating historical usage reports according to the new method, as occurred when the Sourceforge statistics server and system was redesigned, in both 2007 and 2010.[2]

Unreliability of measures poses a threat to validity in two ways. First, it is a threat to statistical conclusion validity because measurement error undermines the ability to accurately assess covariation. Shadish et al. (2001, p. 45) draw on literature to show that unreliability of measures always “weakens the relationship between two variables” and has unpredictable effects on relationships between more than two variables.

Second, these issues can affect internal validity, by undermining the extent to which causality can be inferred from covariance. With digital trace data, where the information system is the *de facto* data collection instrument, there is a risk of mistaking a change in instrumentation, as with a change in use, as a real change to the construct of interest, equivalent to a “treatment effect” in the experimental language of Shadish et al. (2001). This issue arises when a system change occurs in a way such that data collected before and after the change are meaningfully different. As discussed above, systems that are run for the benefit of a community and not for research should be expected to evolve considerably over time, as such technological evolution is a natural outcome of sociotechnical interactions.

In summary, connecting the information system to digital trace data raises issues of reliability that can, in turn, constitute threats to validity. Researchers need to attempt to understand the sources and distributions of such errors and their impact on their chosen measures; one cannot simply assume that errors like these will not be important. To understand the likely errors, intimate knowledge of the online community system and its quirks is ideal. Unfortunately, the system details needed to assess instrumentation reliability are rarely public and often hard to obtain even for participants in the community, who often are not privy to system administration details. Researchers with personal connections who are running the servers or who are otherwise in a position to acquire this information, such as through interviews, have an advantage in establishing the reliability of their measurements. Another option is to undertake small test actions to closely observe how these are recorded by the system. Finally, authors ought to consider the literature on SNA robustness, which will help assess whether their measures are sensitive to particular issues experienced (e.g., Franz, 1998). Reviewers should ask authors to demonstrate knowledge of how the information system affected their data collection and interpretation.

2.2. Aligning Digital Trace Data and Nodes & Links

Any network is, by definition, made up of nodes (vertices, points) and links (ties, relationships, edges). Thus, an important part of the chain of reasoning are the decisions that a researcher makes regarding the nature of both nodes and links. In *Social* Network Analysis (emphasis on Social), nodes are almost always people, although at different levels of analysis they might be individuals, groups, or organizations. Related forms of network analysis, such as Dynamic Network Analysis (Krackhardt &

[1] “Old Wikipedia backups Discovered” http://lists.wikimedia.org/pipermail/wiki-research-l/2010-December/001282.html
[2] http://sourceforge.net/apps/trac/sourceforge/ticket/16511#comment:1

Carley, 1998) and analysis grounded in Actor Network Theory (Latour, 2005) or Socio-technical congruence (Cataldo, Herbsleb, & Carley, 2009) posit a role for nodes representing entities other than people, such as artifacts, tasks, or facts. Kane and Alavi (2008) argue that SNA research in IS would benefit from an approach that includes these multiple kinds of nodes. This perspective specifically includes systems as actors, demonstrating their approach through a study of system use in a healthcare setting that draws on the idea of “indirect system use” through interaction of non-system users with system users.

Table 2. Reliability and System Generated Data	
Decision	Can the system records be taken at face value as accurate and complete? Has the system changed the manner in which it records actions?
Validity issue/type	The information system is the data collection tool and its interpretation is measurement; unreliable measurement threatens both internal and statistical conclusion validity.
Cause	Systems are designed and maintained to serve a purpose other than research; measurement validity is not a requirement.
Examples	Silent truncation of data in Sourceforge and Wikipedia dumps.
Recommendations	• Gain intimate knowledge of the system, through interviews and participation. • Make and track “test” postings, to witness how the system records actions. • Actively inquire about system changes and database purges. • Examine literature on SNA robustness for your intended measure.

Perhaps because they are relatively familiar objects and more or less fixed over time, the conceptual definition of nodes seems to create fewer problems than the conceptual definition of links, leading us to focus on the latter. Below we highlight validity issues stemming from three decisions to be made about links: their type and number, their intensity, and the ontological status of a missing link.

Issue 3: Choosing Multiple or Single Link Types

A key conceptual decision that researchers must make is whether their networks comprise one or multiple different kinds of links between nodes. Borgatti, Mehra, Brass, and Labianca (2009) examine the differences between SNA research as carried out in the social sciences and burgeoning work using similar techniques in the natural sciences, physics in particular. They make the point that social scientists using SNA are usually interested in multiplex links and their interrelationship; as they say, “social scientists typically distinguish among different kinds of dyadic links both analytically and theoretically” (p. 893). These different types of links include similarities (such as location or membership), social relations (such as kinship), interactions (such as communication or sex) and flows (such as flow of information or beliefs). Survey elicitation, sometimes combined with archival data, can be crafted to measure such multiplex links.

Borgatti et al. contrast the multiplex approach with research that has focused on creating massive networks derived from trace data and analyzing their mathematical properties (e.g., their similarity to networks created by processes such as preferential attachment or randomly linked networks). In these networks, there is generally only one kind of link, e.g., a hyperlink between web pages that can be used to derive the structure of the web.

In general, researchers in the IS literature seem to have followed Borgatti and colleagues’ second path, most often constructing networks that include only a single kind of relationship, such as "replied to" interaction (e.g., Wasko & Faraj, 2005). Some studies do utilize multiple sources to draw their networks (e.g., Wagstrom, Herbsleb, & Carley, 2005) but, nonetheless, eventually draw their networks with only a single relationship. A rare exception is the work of Kazienko, Musial, and Kajdanowicz (2008), who studied the photo sharing site Flickr using different kinds of activity such as tagging others’ photos, applying the same tag to a photo, and building contact lists. Eventually, they outline “nine separate layers in one multi-relational social network,” and go on to compare structures in different layers. They do not, however, make strong theoretical arguments that there are separate constructs measured by the different layers, as is more common in sociological applications of SNA (Borgatti et al., 2009).

In summary, IS research studies using SNA have tended to use system-generated data to construct networks of a single link type. This approach contrasts sharply with traditional sociological SNA methods that tend to utilize surveys and interviews, together with some observation, and often collect multiplex relationships. In this sense IS research drawing on SNA is closer to the network research undertaken in physics (e.g., Ebel & Mielsch, 2002; Kossinets & Watts, 2006), than it is to network analysis in sociology (Borgatti et al., 2009). This is true even though the research questions considered in IS typically bear greater similarity to those in sociology than they do to physicists' interest in the topological classification of massive networks and their variation from randomness. While it may be theoretically appropriate to use only single-link types, this is an important decision that needs to be argued from theory and not made merely for convenience.

Table 3. Multiple or Single Link Types	
Decision	Will links be of a single type, or are multiple link types important?
Cause	Found data may only capture a single type of interaction.
Validity type	Construct validity
Examples	Wasko and Faraj (2005)
Recommendations	• Be critical and conservative in assumptions about what links represent. • Triangulate with multiple measures of links (e.g., Wagstrom et al., 2005) and examine consistency.

Issue 4: Defining a Link (Intensity and Dichotomization)

The logical link between data and nodes/links requires researchers to decide what pattern of events constitutes a link and whether that link is binary or valued by its intensity. The intensity issue turns on the argument that the strength of ties affects the nature of interactions between individuals (Granovetter, 1973). Research on SNA in offline contexts has approached this issue by including survey questions on both different types of relationships (friendship, advice, authority) and their respective strengths, allowing participants to translate their memory and interpretation of patterns of past interactions into diverse measures (Marsden, 1990).

Direct interaction data from digital traces would seem to provide useful evidence on interaction intensity, since a count of multiple messages exchanged over time (or other quantifiable link characteristics, like the rate of message exchange or the volume of text in the messages) can be used to indicate varying intensities of interaction between actors by creating weighted networks. However, the decision to operationalize a theoretical relationship based on such data is an inference subject to threats to construct validity. Accordingly, the researcher must carefully use contextual information to guide the selection and interpretation of measures of intensity.

There are a number of techniques for incorporating intensity data in the measurement of a link. One approach is unit weighting, which increases the weight, or value assigned to each link, by a fixed unit for each message between a pair in the network sample. This approach is generally seen in association networks, in which weights represent counts of behaviors, such as an individual editor's changes to specific articles (Kane, 2009). Node strength is also an option for evaluating centrality with this edge weighting method (Valverde, Theraulaz, Gautrais, Fourcassie, & Sole, 2006), indicating the volume of activity in dyadic pairs. Analysis of longitudinal data may apply a time-based decay (Wiggins et al., 2008) to give greater weight to more recent interactions. Most importantly, however, the rationale for these decisions should be presented to demonstrate that the choices made are sensible in terms of the theoretical process held to be occurring.

Complicating this issue, relatively few SNA techniques are intended for use with weighted networks (see Opsahl and Panzarasa (2009) for a summary). Most measures, including all commonly used centralization metrics, assume dichotomous relationships. This assumption is quite appropriate in the design context of limited computational power applied to analyzing networks built on designed surveys that yield abstract relationships of roughly equal strength, as opposed to highly variable interaction-based links from trace data.

As few robust techniques utilize edge weights, the usual analysis approach calls for dichotomizing the networks based on threshold criteria (e.g., only including links that represent more than five interactions). However, dichotomization is a potential source of threats to construct validity that ought to be explicitly addressed. First, dichotomization involves throwing away much of the available source data. Second, dichotomization requires selecting threshold criteria, which can be sensitive to such factors as the size of the data sample. As a result, careful analysis is also needed to determine appropriate theoretical selection criteria for setting thresholds. Finally, dichotomization assumes that the theoretical construct of interest is, in fact, binary, as opposed to continuous. Alternately, rather than treating low levels of interaction as a lack of evidence for a relationship, it may be more appropriate to treat high and low levels of interaction frequency as indicative of different types of relationships, as in Granovetter's (1973) theory of weak and strong ties. It is worth considering, for example, whether links of very different intensities (e.g., one vs. hundreds of exchanged emails) represent qualitatively different kinds of connections. All these issues must be argued on the basis of how best to operationalize a specific construct in the context of an overall theory.

For these reasons, researchers ought to be quite explicit about their dichotomization decisions and should avoid a common pattern of describing the collection of valued data that is then dichotomized for the calculation of the network measure without describing the dichotomization criteria. Unfortunately, decisions about dichotomization are usually acknowledged only in passing or mentioned as a limitation at the end of papers (e.g., Ahuja & Carley, 1999; Crowston & Howison, 2005; Wagstrom et al., 2005), a strategy that confuses the reader as to whether the data collected was, in fact, used, and does not adequately address the validity issues mentioned above. When the interpretations of participants' own understandings of the importance and meaning of past patterns of interactions is not available, the threshold point at which a pattern of interactions (such as count, recency, multiple channels or even content) is sufficient for the inference of the strength or quality of a relationship becomes a key conceptual decision with clear construct validity implications that ought to be argued and explored just as any other issue of construct validity.

Table 4. Link Intensity	
Decision	Should links be binary; if so what is a valid threshold? If not, how should the link value be related to record counts (linear, exponential, through recency?)
Cause	Trace data offers natural counts for intensity, yet these may not match the content of the construct.
Validity type	Construct validity
Examples	Crowston and Howison (2003); Wagstrom et al. (2005); Wiggins et al. (2008)
Recommendations	• Argue intensity decisions, especially dichotomization, with reference to the theoretical context. • Consult Opsahl and Panzarasa (2009) and the TNET R package (http://opsahl.co.uk/tnet/) for measures that utilize intensity.

Issue 5: Defining a Non-Link

The choice of when to assess that a link exists is also a choice of when to assess that a link does not exist. In many theories, the absence of a link is as meaningful as its presence. For analyses drawing on the notion of brokerage or "structural holes" (Burt, 1992), for example, it is fundamental to understand where information cannot travel, since this identifies privileged routes (a broker is one who is *uniquely* linked to a portion of the network and, therefore, able to control access or information flow; a structural hole is one of the missing potential links between groups that could be strategically filled). The construct validity of such measurements depends on the validity of the inference that the network is one in which information flows along the identified links, but just as importantly, that information cannot flow where links have not been identified.

Similarly, the meaning of non-links is important to understanding the construct of information sharing, important in innovation, diffusion, and contribution (e.g., Brynjolfsson et al., 1994). Information sharing can be studied from a network perspective by measuring the network of individuals linked through their communication activities. Given a valid information-sharing network, SNA summary measures can provide insight into the processes of information sharing by identifying key individuals and providing measures for comparison of different groups. For example, high betweenness centrality indicates which individuals are on the shortest path between many others and, therefore, positioned to affect the flow of information through the network. Likewise, network diameter indicates the maximum number of links through which information must travel in order to be transmitted between an average pair of individuals, suggesting how quickly a group may spread new information. Again, the validity of such measurements depends on the assumption that the absence of a link means information cannot flow.

Traditionally recommended SNA techniques, such as survey responses to name generators, implicitly provide non-occurrence data. Asking survey respondents to indicate all of the people with whom they interact from a list creates valid grounds for inferring that those not indicated are not interacted with (at least not sufficiently for the respondent to infer a relationship). However, to connect digital trace data to nodes and links requires the researchers themselves to make this step and to demonstrate that they have done so with sufficient validity. In some cases, the absence of any events suggesting a link may be an appropriate indicator of the absence of that link, but this assumption is not always justifiable (see Borgatti, Carley, & Krackhardt, 2006 for a detailed discussion). As a result, it is incumbent upon the researcher to be clear about the ontological implication of the absence of evidence regarding a link. Just as researchers must argue that their inference of a link is valid, they must also argue that their inference of the absence of a link is valid.

When analyzing face-to-face networks, inference from missing evidence to non-links is bolstered by physical aspects of the world, such as the limited range and impermanence of audio and the real-time feedback between speaker and listener; evidence of speaking to another is both evidence that the other heard and evidence that others not present did not hear (at least not through this event). Such an assumption may also be valid for interaction via some ICT, as when emails are exchanged directly from senders to a short list of recipients listed in the message (i.e., non-broadcast email), especially when those recipients reply, indicating that they had, in fact, received the message.

On the other hand, trace data often includes listservs or other broadcast forums, especially in online communities. In most listservs, all emails are archived and made available to all community members, and even to the general public (Grippa, Zilli, Laubacher, & Gloor, 2006). When email communications occur via a listserv, whether archived publicly or not, the data provides weak evidence regarding information flow and control. In particular, it is impossible to argue the meaningfulness of measures based on information control, such as betweenness or closeness, as measures of importance, because in this case there is no such mediation. Calculations such as the diameter of a reply-to network are similarly meaningless for understanding information flow: If information is broadcast on a mailing list, it potentially reaches all group members at once. Unfortunately, a lack of consideration of the properties of the medium is disturbingly common in IS research, and rarely addressed (e.g., Bird, Gourley, Devanbu, Gertz, & Swaminathan, 2006; Concas, Lisci, Pinna, Porruvecchio, & Uras, 2008; Wu, Goh, & Tang, 2007). Truly grappling with information flow in discussion lists would require an understanding of readership behaviors. Unfortunately, very little work has directly examined readership, since it usually leaves no trace data; notable exceptions are Lakhani and von Hippel (2003), Yeow, Johnson and Faraj (2006) and Goggins, Galyen, and Laffey (2010).

Consideration of the meaning of non-links suggests validity concerns regarding a common analysis strategy with data from listservs, namely the analysis of reply-to links. As message recipients are not specifically named in mailing list data, researchers often examine instead the structure created by message responses (e.g., Crowston & Howison, 2005; Wasko & Faraj, 2005; Wu et al., 2007). A network can be constructed by creating links between message authors at the message level, linking A to B if B replies to a message posted by A. Or the network can be constructed even more indirectly, at the level of the reply thread, by creating a link between all participants in a given email reply thread (as in Concas et al., 2008). Unfortunately, few researchers have been adequately explicit about what construct such a network represents (i.e., what the presence vs. the absence of a reply means conceptually). It should be clear, at least, that response structure is not a valid measure of information flow: While those who reply to a message have (most likely) read it, non-response does not indicate that other members have not. Messages posted to an email list may be read by only the people who reply in a given thread, by every member of the list, or, more likely, by some unknown proportion of the subscribers (Howison, Inoue, & Crowston, 2006) and possibly even non-community members accessing a listserv archive.

Our point is not to argue that networks constructed from broadcast reply-to trace data cannot be useful or ought not to be explored. Such network measures might, in fact, provide some very interesting insights, such as who or what prompts another to reply in public, or allow researchers to make non-information flow arguments based on, for example, the signaling effect of having been replied to (i.e., by providing an argument for the interpretation of a reply in a broadcast context vs. an absence of a reply). Our point is merely that the researcher should make an argument as to the meaning of such links explicit. More generally, researchers should take as much care to argue that the identification of a missing link is valid as they do to argue the presence of a link.

Table 5. Missing Links	
Decision	Are missing links theoretically important? If so does the absence of a positive link validly provide evidence for the absence of that link?
Cause	Trace data are the result of action but may not provide evidence of inaction for some constructs.
Validity Type	Construct validity
Examples	Crowston and Howison (2005); Wasko and Faraj (2005); Wu et al. (2007)
Recommendations	Understand the theoretical significance of missing links; explore whether unrecorded actions (such as reading) need to be considered.

2.3. Aligning Node & Link and Network

The next set of issues concerns the logical connection between appropriate definitions of nodes and links based on well-understood digital trace data and construction of a network. Making this connection can seem deceptively simple but can pose significant threats to validity. The key challenge stems from trace data as longitudinal data: Events occur at particular points in time, and, thus, multiple events must be aggregated to construct a network.

In SNA based on surveys, data are collected at a particular point in time, but as they are based on recollections, by nature, they measure impressions up to that point in time. Such an approach is appropriate to measure relatively stable links. Indeed, many sociologists prefer survey data for exactly this reason: They capture participants' understanding of the social relationships in general that is typically the construct of interest, rather than the interactions at a particular moment in time, which may or may not be representative of the network (Marsden, 1990).

In contrast, trace data are records of events that take place at particular points in time, and those events can be quite sporadic (e.g., a series of email messages sent from person to person). Data representing associations may also be available longitudinally, such as records of members joining, leaving, or participating in groups (e.g., editing a wiki page at a particular point in time).

The longitudinal and episodic nature of trace data offers both opportunities and threats to validity. On the one hand, longitudinal data can be very valuable for testing causal theories. For example, Hahn et al. (2008) studied the effect of previous working relationships on later decisions about which open source software project to join. Other researchers have taken advantage of the temporal nature of the data to investigate network dynamics, e.g., by drawing networks for consecutive time periods, thereby producing time series of network statistics and analyzing the trends (e.g., Christley & Madey, 2007; Falkowski, Barth, & and Spiliopoulou, 2008; Howison et al., 2006; Long & Siau, 2007). Researchers have also explored visualization techniques for longitudinal social networks (Moody, McFarland, & Bender-deMoll, 2005), and more specifically, for handling the fine-grained temporality of online discussion data (Trier, 2008). On the other hand, longitudinal data must be aggregated to build a network structure (Trier, 2008), collapsing a series of events over time. The extended period of data collection and the necessary aggregation process have implications for the construct validity of the resulting network measures (Howison et al., 2006). We examine two in detail below: temporal aggregation and temporal mismatch.

Issue 6: Temporal Aggregation

A particularly pernicious issue arises when creating a network by aggregating links that occur at different points in time. For example, consider a study of information sharing using point-to-point communication links, where A sends a message to B and, later, B sends a message to C (see Figure 2). If the messages are sent in this order, it is possible for A's information to reach C, but not if the messages occur in the opposite order (in the absence of other messages, as we discuss below). Similarly, in the case of an association network, if two individuals are members of a group at the same time, there is a possibility of some kind of influence process (such as learning of best practices), but if their memberships do not overlap in time, the influence can be in one direction at best (e.g., Kane, 2009).

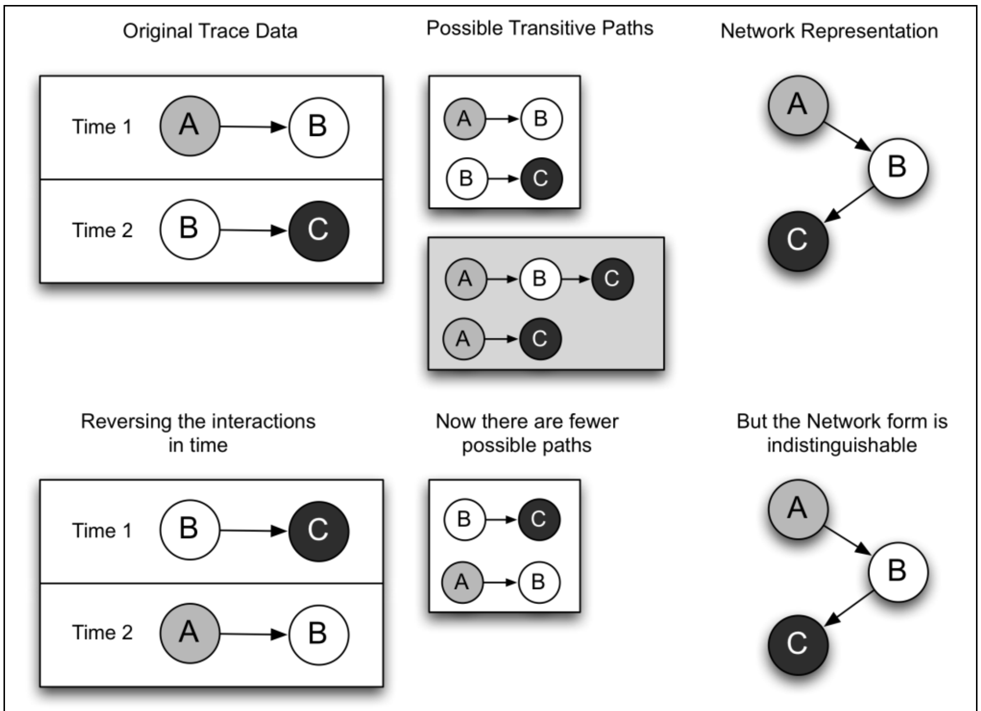

Figure 2. The Implications of Collapsing Flow Networks Built from Trace Data Over Time; Note the Indistinguishable Network Representations for the Different Sets of Possible Paths

Aggregating links across time to form a single cumulative network will suppress these nuances, potentially leading to invalid conclusions. When working with flow networks, at least, even employing a directed graph representation can introduce paths not possible in the original data, as demonstrated in Figure 2, below. Since the logic of many common network summary measures is based on paths through the data (see section 8 below), the introduction of impossible paths due to temporal aggregation is a clear threat to construct validity. (It might be less problematic in networks that are not based on the logic of flow, see Discussion, below). Avoiding this issue entirely can be difficult; aggregation is required to perform network analysis using digital trace data.

Two techniques are available to deal with the issue. The first approach is to represent the "network" as a set of actual sequential paths through nodes, rather than a traditional network, and then to analyze it appropriately, an approach demonstrated by Brynjolfsson et al. (1994).

A second approach is to follow the argument of Nia, Bird, Devanbu, and Filkov (2010) (who respond to a working version of this paper). They call this issue "transitive faults" and demonstrate two approaches to exploring its impact. Their arguments are empirical; they make the case that this issue is not problematic for their specific data, rather than in general, however their approach could be followed to confirm this for any set of specific data.

Their first technique is to develop upper and lower bounds on the quantity of "transitive faults" created by different time windows (measured by Spearman rank correlations between the results for each sized time window. Such bounds are an excellent approach to arguing to show that the issue does not significantly affect results for particular data and a particular research question.[3]

Nia et al.'s (2010) second technique is to use a simulation of network growth to "fill in" the missing data and then show that the measures of interest have reasonable correlations, whether created with the original data or the simulated data. This second technique relies on knowing an appropriate simulation of behavior leading to the network and understanding that the data collected is not complete (see Issue 9, below).

Table 6. Temporal Aggregation	
Decision	Does the order in which events happened matter? Will aggregation introduce spurious or empirically impossible links?
Cause	Trace data capture evidence of dyadic links; a network must be an aggregation of such links. Aggregating directed links introduces spurious links.
Validity Type	Construct validity
Examples	Howison et al. (2006); Kane (2009).
Recommendations	• If the links are directed, consider working directly with network paths, rather than collapsing to a regular network (Brynjolfsson et al., 1994). • Explore and demonstrate upper and lower bounds on this problem for your data and measure, arguing that even if the measure is affected to the extent of the upper bound, the results will still support the argument made in the paper. See Nia et al. (2010).

2.4. Aligning Network and Network Measures

A common task in the analysis of a given network is to compute various measures of the network. For example, as noted above, in studies of influence, betweenness centrality might be computed to determine which individuals are positioned to affect the flow of information through the network. However, the longitudinal nature of the trace data raises validity issues in this task. In addition, differences between digital trace data and more typical SNA data are reflected in the potential mismatch of SNA tools used for such calculations to trace data.

Issue 7: Temporal Mismatch

A decision about the time period over which to construct a network is simultaneously a decision about the period of time for which measures derived from that network will be measured. An issue of construct validity from aggregation comes from a potential mismatch between the stability of the construct of interest as compared to the degree of aggregation of the data. The particular construct measured as a network link may be conceptualized as being stable (e.g., long-term friendship ties) or dynamic (e.g., high school dating ties), meaning that the network structure potentially changes and evolves over time (see Huisman & Snijders, 2003; Leskovec, Kleinberg, & Faloutsos, 2005). Of course, stability is relative, depending on the time scale involved. Social relations may be stable for months or years but perhaps not for decades.

[3] While we endorse the overall methodological approach of Nia et al. (2010), their specific application seems problematic since they limit their analysis to the top 10 percent of participants by message count. This makes it much more likely that, as time windows expand, an exchange will eventually be found that resolves the transitive fault. For some research questions, such as those concerned with diverse sources of knowledge from the periphery, this decision would undermine the usefulness of the technique. In general, however, seeking and showing upper and lower bounds for the impact of this issue is an excellent approach.

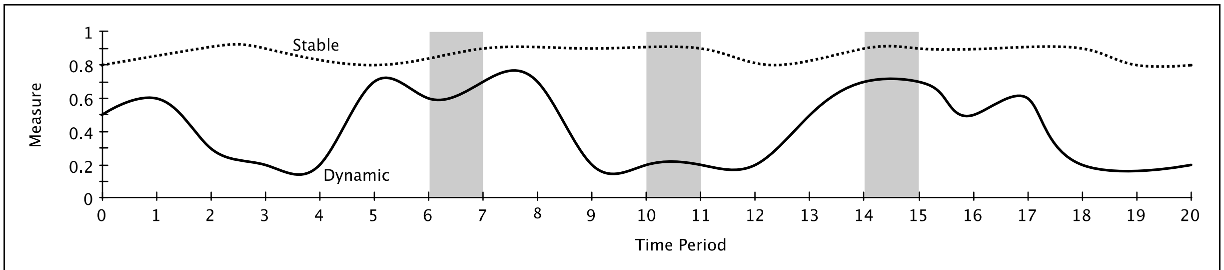

Figure 3. Validity Issues Deriving from Mismatch of Aggregation and Construct Stability

The combination of these two characteristics of network data—temporality and construct stability—may threaten the construct validity of network measures created when aggregating digital trace data across time (Braha & Bar-Yam, 2006). Figure 3 shows illustrative data; the top line (dotted) shows a relatively stable construct, the lower line (solid) shows a construct that varies considerably over time. The sections marked in grey show potential snapshots.

The top line in the figure shows a case with no significant concerns: The constructs of interest are stable, so the aggregation of interactions in the form of snapshots or aggregated measures will yield similar results. For example, networks of familial relationships will show more-or-less the same links in both snapshot and aggregated representations, with the exceptions of the addition or subtraction of actors over time due to birth, death, marriage, and divorce.

However, if the constructs are less stable (the bottom line), then a snapshot will measure the network configuration only at that point in time, assuming that the snapshot size and the construct's stability are appropriately matched. In Figure 3, snapshots taken in the three grey areas approximate reasonably well the up-and-down cycle of the measure. Although the network structure may be different at other points in time, the measure may still provide useful insights into social processes. Concerns would arise, however, if data were only taken at the first and third snapshot, since the result would be an invalidly high and consistent measure.

The case of aggregating data about unstable constructs is the most problematic. There are two issues here. The first issue is relatively well known: The average of a network measure taken over time will smooth out important variance. The second issue is less well understood and is more clearly the result of aggregating events and drawing networks: The resulting network may have very different structural properties depending on how events are aggregated.

For example, Howison et al. (2006) examined centrality in open source development teams initially by aggregating interaction data across the life of projects. They were surprised to discover that while some projects had only a few or just one highly central developer, as hypothesized, other projects had many apparently central actors, suggesting a relatively decentralized team structure. However, when they examined the data dynamically, they discovered that a much greater number of the projects exhibited a high degree of centralization at any point in time, but in some, the most central actor changed from time to time. In other words, the role of lead developer was unstable in some projects. It was only when this series of centralized networks were aggregated that the resulting network appeared to have multiple central nodes, and, thus, appeared to be decentralized, as illustrated in Figure 4. The choice to measure centralization on an aggregated network assumed that this construct was relatively stable, leading to invalid conclusions about the projects.

This concern is primarily an issue of construct validity: What period of aggregation leads to a (approximately, usefully) "correct" understanding of the network? Another way to think about this would be to ask, "Over what period of time does the network process of interest play out?" or, depending on one's stance on how networks influence action, "Over what period of time does network structure come to influence action, such that the actions validly approximate the network that influenced them?" While these are primarily issues of construct validity, they can also be thought of as issues of measurement error and, thus, relevant to internal validity.

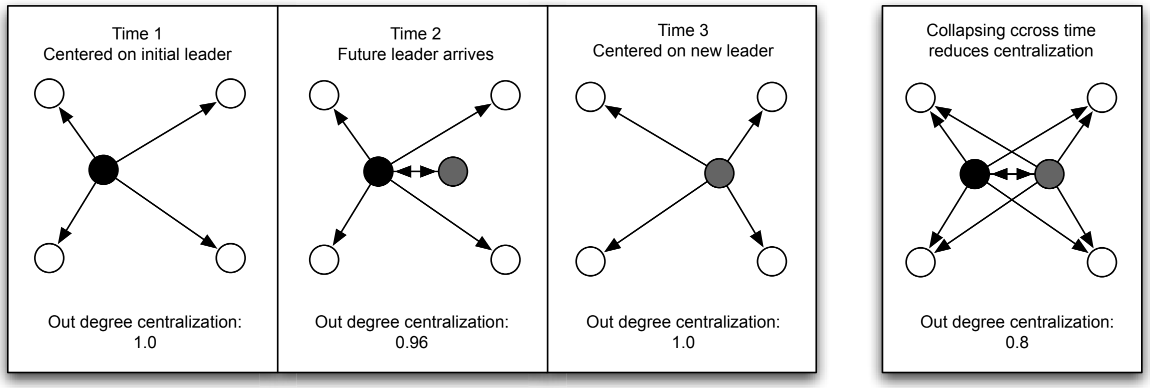

Figure 4. Aggregation of Data with Unstable Construct (Here Leadership) Artificially Decreases Centralization (Adapted from Howison et al., 2006)

One approach to dealing with this issue, especially for dynamic concepts, is to vary time windows to locate a periodization over which one's construct is more reliable. Olson and Carley (2011) describe a method (using Cohen's Kappa and information loss) to explore the reliability of measures over time and identify window sizes in which measures are most reliable. Such methods, in combination with arguments from theory about the likely length of time over which the network process of interest plays out, would help to establish that research has avoided this threat to validity.

Table 7. Temporal Mismatch

Decision	Over what period will events be aggregated to form networks (and thus measure network concepts)?
Validity Issue/Type	• A dynamic construct may invalidly appear static if measured with long aggregated networks; an otherwise stable construct may invalidly appear dynamic if measured on too short a time scale. • Aggregation over long time scales may produce networks with different structural properties than the network experienced by participants.
Cause	Trace data capture evidence of dyadic links; a network must be an aggregation of such links and, thus, occur over some time period. Constructs may influence action in ways that are only visible over some particular time scale.
Examples	Howison et al., 2006
Recommendations	• Assess theoretical stability of construct and likely time scale. • Conduct sensitivity analyses to assess the effect of different periods of aggregation, using agreement statistics to measure impact. See Olson and Carley (2011). • See Braha and Bar-Yam (2006).

Issue 8: Network Tool Effects

Social Network Analysis is greatly facilitated by a wealth of software tools that implement a wide range of algorithms. Popular tools include UCINet (Borgatti, Everett, & Freeman, 2002), Pajek (de Nooy, Mrvar, & Batagelj, 2005), the SNA package for R (Butts, 2008), and NodeXL (Hansen, Shneiderman, & Smith, 2010). In general, these tools are excellent in terms of validity: They help researchers avoid errors that might stem from re-implementation of algorithms and provide consistency and reproducibility across different researchers.

Nonetheless, the convenience these tools provide can also mask threats to validity in their use. First, programs use subtle variations of algorithms and slightly different names for the same algorithm, potentially leading to confusion and misinterpretation of results.

Second, tools make the (reasonable) assumption that the data provided are appropriate for the calculation requested. Just as with more familiar assumptions in other statistical techniques, such as cell size for ANOVAs or normality for some types of regression, a tool may or may not highlight these assumptions. For SNA, it is rare for the tools to do so. For example, some very common algorithms (such as degree centrality/centralization) work properly only with dichotomous data (binary links without weighting). Tools may, therefore, assume that the user intends that the data be dichotomized. If valued data are presented to such routines, the tool may silently introduce dichotomization at strength >= 1, a decision that can threaten validity (see Issue 4, above), or may simply carry out the calculations with inappropriate values.

For example, while the definition of degree is operationalized by *counting* the number of links, the network degree centralization function in the SNA package in R *sums* the values in the matrix by default. If the link values are binary (unweighted), this is an equivalent approach, but if they are weighted, then the function silently performs a weighted centralization function. This is a much less commonly understood and interpretable measure (see Opsahl, Agneessens, and Skvoretz (2010) for a discussion of this and alternative measures). If the link values are not explicitly ignored, the software produces a result for degree centralization that is quite possibly not what the user intended.

Finally, and most subtly, algorithms embedded in tools may make assumptions about the nature of the data, assumptions that interact with issues discussed above to produce threats to validity. For example, a class of algorithms, including eigenvector centrality, is justified through logic that treats the network as a topology and constructs all possible paths (or an infinite length random walk across those paths) from the network representation. Similarly, closeness, betweenness, and many grouping algorithms make assumptions that long paths are relevant and possible. The computation can, thus, invoke paths that may not be justified by the theory in use, creating validity issues (see Issue 6, above, and Issue 9, below). The design of network algorithms is a situated practice, drawing on particular types of networks and network processes; a mismatch between their internal logic and network characteristics can introduce validity issues.

In short, just as with any statistical package, the convenience of tools does not eliminate the responsibility of the authors and reviewers to be sure that they are used appropriately. Tool authors are generally careful to provide references that describe their algorithms in detail. Authors should find such references and examine the assumptions of the algorithms. Authors should build confidence that they are using the tools correctly, for example, by manually calculating a measure for a small prototype network and comparing it to the tool's answer. An alternative is to calculate the same measure with multiple tools and carefully understand the reasons for any differences. Authors should be prepared to provide complete step by step descriptions of their tool use (or, ideally, scripts) to help reviewers and readers judge its validity and to enable others to replicate their method (such descriptions are known as research protocols in the natural sciences, and typically published as online addenda.) Careful consideration of validity issues stemming from tool use will improve the validity of network analysis.

Table 8. Network Tools	
Decision	What SNA tool/software will be used? Is the algorithm cited? What assumptions about the data is the tool making?
Validity Issue/Type	Multiple
Cause	Software tools perform much of the "heavy-lifting" in network analysis, but algorithms may be influenced by default settings or subroutines that encode hidden assumptions (e.g., silently dichotomizing valued links).
Examples	Errors such as these are not visible in papers and can't be checked unless all data and analysis scripts are provided. We encountered these issues in our own research and confirmed that other users were not aware of these issues.
Recommendations	• Build confidence through manual calculation, tool triangulation, and known outcome tests. • Methodologists and tool builders: Make the assumptions contained in algorithms and tools explicit.

2.5. Aligning Measure and Construct

Measuring a theoretical construct using network data is, of course, the reason to undertake the work in the first place. This alignment is between the concrete and the abstract; the argument that a network metric is an appropriate measure of a construct ought to be carefully considered and its validity explicitly argued. In the validity framework of Cook and Campbell, this issue very closely matches construct validity. In this sense, a network measure is an operationalization of a construct, and general recommendations for demonstrating construct validity apply, including face validity, congruent validity, and discriminant validity.

Face validity. Face validity is perhaps the simplest yet most overlooked aspect of validity. An excellent candidate for showing it is to provide concrete narrative examples of the hypothesized process drawn from the dataset. As discussed above in Issue 1, the digital trace data often provide rich data as a basis for such narratives, which might be effectively complemented by interviews. Even a single clear case of a hypothesized process, together with an argument that the proposed networks and measures validly measure it, can go a long way toward exposing validity concerns. Once exposed, these concerns can be dealt with explicitly, enhancing the usefulness of the approach. If authors cannot describe a single clear case from their dataset, skepticism is warranted.

Congruent and discriminant validity A useful strategy for demonstrating the validity of any measure is to show congruence between that measure and other, independent, measures of that construct. This simultaneously avoids mono-method bias and argues for the validity of a proposed measurement technique. For example, if one intends to use network centrality as a measure of leadership, then a demonstration that this measure has adequate agreement with other appropriate measures—such as lists of those nominated by a community as leaders on a web homepage, or interview or survey results—would be useful. If such agreement is not forthcoming, then the authors ought to be able to explain why their measure is different yet still appropriate. Similarly, it is appropriate to show that one's measure is relatively unrelated to conceptually dissimilar constructs, such as showing that leadership is distinct from simple counts of activity (unless one's theory of leadership directly involves counts of activity).

Issue 9: Data Completeness and Inference

The basic structure of many social network theories hypothesizes an unobservable social relationship (the construct of interest) that leads to various kinds of interactions that can be observed, for example: a friendship relationship that leads to observable conversations, or an information sharing relationship that leads to observable questions and answers. Thus, the existence of the relationship is inferred from the observed interactions. Furthermore, in offline observational data collection, researchers expect to observe only a fraction of the interactions between individuals: There are understood to be many more interactions than periodic or partial observation can measure. Therefore, the observation of a specific interaction that is indicative of a relationship can be assumed to indicate the presence of many similar unobserved interactions. The logic of these inferences is as shown at the top of Figure 5, below.

In other words, the interactions among members of a community can be thought of as a population generated by the social relationships from which the particular observations (or reported links) are somehow sampled, allowing the application of inferential logic to make claims about this population of interactions and the relationships for which they may provide evidence. For example, in studying knowledge sharing, the analyst might observe a set of spoke-to interactions between two participants and interpret this as evidence for the existence of a relationship of interest, inferring the likely existence of other, unobserved, spoke-to interactions that could provide channels for information transmission, influence, or other network processes. In many face-to-face groups, it might further be assumed that the intensity of interactions is roughly comparable, and that all interactions are at least potentially two-way (i.e., an assumption about the likely distribution of interactions in the population of interactions).This again facilitates inferences about the population from the sampled interactions.

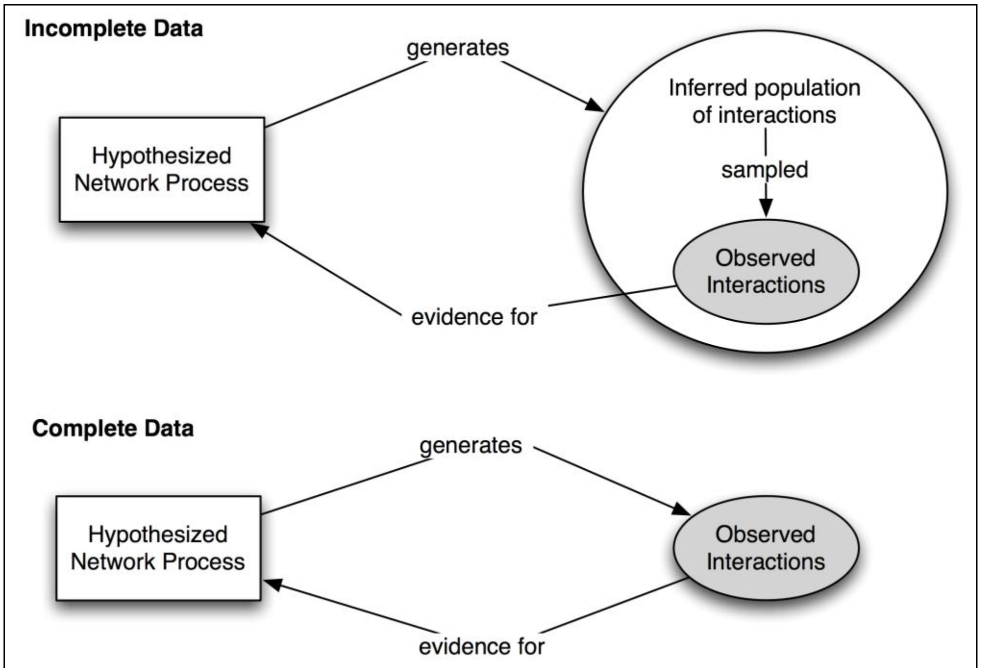

Figure 5. Comparison of Inferential Logic Applied to Partial and Complete Data

In contrast, with digital trace data where the Information System archives every interaction, and when there is good reason to believe that the group only interacts via this platform, the data provide complete evidence of interactions, a census rather than a sample of interactions, as shown at the bottom of Figure 5. This situation is actually quite common in studies of online communities, many of which only exist virtually. In this situation, the hypothesized relationship continues to generate events, but rather than this producing an unknown population from which the observations are a sample, the researcher can access the full population of events that did, in fact, occur.

On the one hand, the completeness of the data is a good thing, as it allows more definite conclusions to be drawn based upon the observed dynamics. Researchers using these data have a rare and enviable degree of certainty that the data are comprehensive. On the other hand, researchers using such data must be wary of the human tendency to infer structure from interactions and assume that evidence based on a set of events is representative of deeper meaning. In the case of trace data, what you see may be all there is. There is no need to postulate that the observed interactions represent a partially hidden pattern of interactions; the pattern, if there is one, is in fact quite explicit.

Furthermore, when data are from the full population, techniques designed to work with samples can give meaningless results. In the Cook and Campbell framework, this situation poses an issue of statistical conclusion validity, albeit one that rarely arises: Researchers can readily acquire sufficiently complete data such that inferential statistics or thinking are no longer necessary or appropriate, and this requires thinking differently about the analysis. In particular, depending on the construct of interest, inappropriate use of inferential logic potentially poses a potential threat to validity in a wide range of analyses (e.g., Aral et al., 2006; Kane, 2009; Merlo et al., 2009; Wasko & Faraj, 2005).

As a concrete example, consider again a study of information sharing behavior. In a face-to-face group, the observation that Person A spoke with Person B in Week 1 of a study might be taken as evidence of a relationship from which the analyst might infer the likely existence of other unobserved communication events, forming a two-way link through which information could travel. The validity of this measurement relies on the inference that if Person A and Person B are observed to speak at some point in time, Person A likely speaks with Person B at other times, generating a population of interactions, as shown at the top of Figure 5. Indeed, this inferential logic is behind the approach of creating a network as shown in Figure 2: Having observed only the second set of interactions, the researcher assumes that the additional interactions in the first set are likely to have occurred at some unobserved point in time, and so implicitly includes these interactions in the measurement.

Contrariwise, if the researcher is reasonably confident of having observed all interactions in the group (the situation at the bottom of Figure 5), this form of inferential reasoning and the conclusions based on it are invalid. Regardless of any relationship that may be suggested by Person A speaking to Person B in Week 1, if the data do not show that the two speak again, then there is no evidence of a two-way information channel; indeed, the data rule it out, at least in the period under observation.

Inappropriate use of inferential logic also poses a threat to some studies using association network data. While association networks are often used to indicate overlapping interests, they are sometimes used in ways that require them to be a proxy for interactions (e.g., Daniel & Diamant, 2008; Grewal, Lilien, & Mallapragada, 2006; Kane, 2009). For example, researchers might use joint membership in a project as a measure of possible knowledge sharing among members. Such an inference is unnecessary, and may, in fact, be invalid if detailed interaction data are available that circumscribes the possible paths or when temporal overlap data regarding membership is available (e.g., Christley & Madey, 2007; Merlo et al., 2009). Brynjolfsson et al. (1994) and Hahn et al. (2008) study interaction paths directly, rather than networks, and so are notable for avoiding this issue.

In summary, interpretations that tacitly or explicitly rely on inferential logic should be considered suspect when it is likely that the data show close to the totality of interactions. Unfortunately, as demonstrated in Figure 2, making this assumption can occur in the very act of drawing the network, where impossible indirect paths are introduced to the network by temporal aggregation. Similarly, as mentioned above, some network algorithms have sampling logic built in because they work by back-constructing a set of all possible paths from a network diagram, only then using the paths to calculate the network measure.

In different contexts, this issue might be less of a problem. First, in some circumstances it might be quite reasonable to assume that the observed events are an incomplete record and that additional interactions occurred, perhaps by unrecorded media such as instant messaging, private email, or face-to-face interactions. Second, even fully complete data for one period do not circumscribe all possible interactions that *could* be generated from a relationship (see Discussion below), so complete data from one temporal period may be considered a sample of all possible interactions and, thus, predictive of future unobserved interactions. Such sampling logic, however, must be argued to be reasonable; there is nothing in the construction of a network that relieves the researcher of that responsibility. Further, some network properties may be robust to certain patterns of missing data, and appropriate with smaller proportions of the network, while others may not be (for a detailed discussion, see Latapy & Magnien, 2008).

Table 9. Data Completeness and Inference	
Decision	Is my data a sample or a census of activity?
Validity Issue/Type	Statistical conclusion validity
Cause	If the data approaches a census, sampling logic may be inappropriate. Sampling logic, realized in some SNA algorithms, may introduce and interpret events known not to have occurred.
Examples	Daniel and Diamant (2008); Grewal et al. (2006); Kane (2009)
Recommendations	• Consider carefully how sampling logic is employed, and argue for its appropriateness. • Consider whether network algorithms introduce events known not to have occurred. • Consider whether associations are valid proxies for interactions (if the association network is being used in such a way). • Consider using methods in Brynjolfsson et al. (1994) and Hahn et al. (2008)

Issue 10: Uncritical Importation of Measure Interpretation

The final link is between measures and interpretation as a theoretical construct. A regrettably common threat to validity arises when researchers import interpretations of measures from previous literature without considering whether the underlying networks (nodes and links) for which these measures and interpretations were developed are conceptually similar to the context and type of data in the present study. While this problem could occur with any study, it appears to be particularly tempting when working with found, rather than designed, data sources. Thus, it is particularly likely to affect work with digital trace data. Importing interpretations of measures based on survey data to networks built from trace data are particularly common and often problematic.

Early work, such as Ahuja and Carley (1999), makes the importation of concepts explicit and considers it critically, outlining findings from offline environments and providing a rationale for their applicability in online contexts, specifically questioning whether the concepts and measures will be appropriate to the new environment. Other works, such as Wu et al. (2007), have been less careful to problematize their adoption of interpretations based on earlier work, instead making claims such as “Past research in social networks has shown that centrality is an important indicator of group performance” and citing as warrant an SNA classic such as Freeman, Roeder, and Mulholland (1980). The truth, or usefulness, of this statement depends on how cohesive the entire chain of reasoning is: The meaning of centrality depends strongly on decisions about nodes, links, and measures (e.g., exclusive channels of communication vs. broadcast communication), all taken in a particular theoretical context. In short, the environment in which the data were generated influences the interpretation of network measures. Unfortunately, many studies are surprisingly vague about the theoretical rationale for the choice of a particular construct and its connection to the data. Many rely on ill-defined notions of general, abstract ties as though any graph structure, however defined, is a valid proxy for the same abstract concepts (i.e., mistaking SNA for a theory rather than an analysis technique).

Researchers and reviewers should be particularly aware of this issue and work to avoid the importation of an interpretation from earlier studies without an explicit argument for its appropriateness in terms of theoretical cohesion among node, links, measure, and construct. It is possible for researchers to hold a considered position that any set of connections, however defined and measured, operate in a usefully similar manner; but if so, they ought to be explicit about this, as it is an extreme position. It is certainly not sufficient to imply that since SNA techniques are being used, importation is *prima facie* valid.

Table 10. Inappropriate Importation of Network Measure Interpretations	
Decision	On what logic are interpretations of networks measures based?
Validity Issue/Type	Construct validity
Cause	Network measures are associated with networks built from particular data and may not be valid outside their original context.
Examples	Wu et al. (2007)
Recommendations	• Understand and explicitly argue for a correspondence between definitions of nodes, links, and network measures based on network processes. • Explicitly argue that an interpretation from earlier SNA studies is appropriate, given your data (e.g., Ahuja and Carley, 1999). • Draw on Borgatti and colleague’s taxonomy of network processes (see Discussion).

3. Discussion: Maintaining Overall Theoretical Cohesion

While we have presented them separately, the issues raised above are, of course, not independent. Researchers employing SNA (with or without digital trace data) have to maintain cohesion among all of these logical links in order to mitigate validity issues. The theory with which the researcher is working is fundamental to this task. In particular, as we argue below, the type of network process entailed by the theory binds together the logical links and brings cohesion to them. This cohesion is the central bulwark against validity issues.

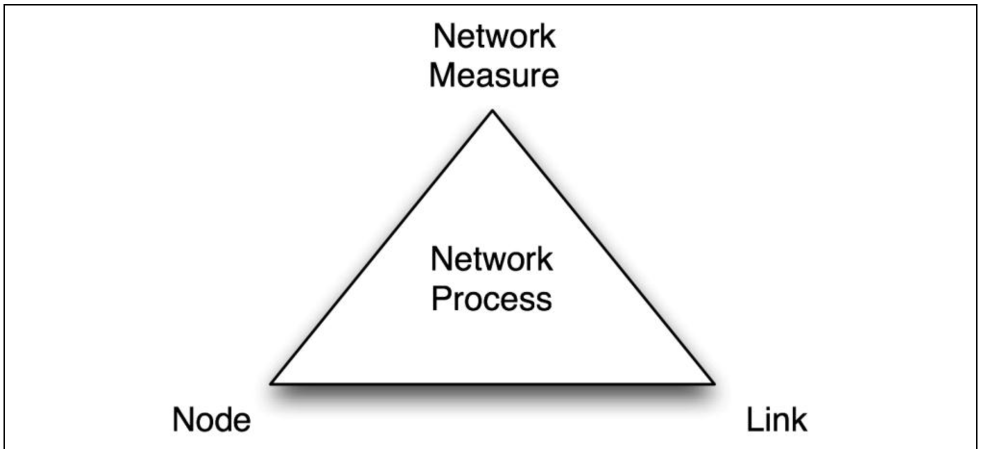

Figure 6. Network Process Provides Theoretical Cohesion to SNA Decisions

Of particular assistance in this endeavor is work by Borgatti and colleagues that builds a taxonomy of tie types and relevant network processes. The first distinction is between structuralist and connectionist perspectives on networks (Borgatti & Foster, 2003) and the second is a taxonomy of types of network processes (Borgatti et al., 2009).

3.1. Structuralist vs. Connectionist Views of Networks

With regard to the first, the structuralist view focuses on ties as a *topology*, while the connectionist perspective sees ties as *instances*. The structuralist view is that the network describes a topology on or through which the phenomena of interest are assumed to occur. The connectionist view is that the links do not form the topology (what could occur), but instead represent the actual events of interest (what occurred). Trace data, as defined in this paper, are inherently closer to the connectionist perspective: they represent *instances*. By contrast, research based on asking people about social relationships (the traditional approach to SNA) is typically structuralist: The surveys attempt to measure structure that, from time to time in some manner, influences events.

These two views require different ways of interpreting the data, but are often confused, leading to validity issues, as described above. Unfortunately, when working with trace data, it seems there is a tendency to take evidence of instances (what was) and transmute that uncritically into evidence of topology (what could be/have been). To avoid this problem, researchers should be clear about whether their theory is a theory about structures or instances. If one's theory requires understanding of structures, but one has evidence of events, then one must reason from the instances to the structures. Such reasoning is not impossible, but it requires an explicit theory of how structures are created by events, and how events create structures. This consideration suggests that relevant theories would be those grounded in structuration or practice theory (e.g., Contractor et al., 2000; Giddens, 1984; Orlikowski, 1996).

The difficulties in linking data about events to evidence of structures underlies many of the issues discussed above, including issues related to deciding between single or multiple link types, coping with intensity, and coping with temporal aggregation and temporal mismatch. This distinction also helps understand why importing interpretations of network measures from earlier work is problematic: if the interpretations are based on measuring evidence of structure (as many are), then their logic breaks down when working with data which are instances.

3.2. Network Mechanisms

Theoretical cohesion can be further improved by a consideration of types of network mechanisms (which may play a role in many different processes). Borgatti et al. (2009) identifies four types of network mechanisms: transmission, adaptation, binding, and exclusion. *Transmission* networks involve the transmission of something between network nodes, *adaptation* (or similarity) networks posit links based on similar experiences of nodes, networks based on the *binding* mechanism result when “social ties can bind nodes together in such a way as to construct a new entity” with its own properties. Finally, an *exclusion* mechanism involves “competitive situations in which one node, by forming a relation with another, excludes a third node.” Network mechanisms are more specific than network processes: for example, influence could be conceptualized as a network process, occurring through multiple mechanisms, including information transmission, similarity binding, and exclusion.

Borgatti (2005) provides detail on transmission mechanisms, the most common type of mechanism considered in IS. These mechanisms involve the transmission of something between network nodes, and can be classified according to whether that thing is thought to move by a copy mechanism (such as ideas) or a move mechanism (such as money), as well as the type of path through the network that the thing follows (e.g., shortest path, random path, or parallel paths). Each mechanism implies different ways of measuring links and different processes occurring over these links; different theories, when carefully considered, involve different mechanisms. Borgatti and colleagues argue that a valid match between mechanism and network construction—which can only come from a strong theoretical understanding—is key to choosing the appropriate measures, as “different measures make implicit assumptions about the manner in which things flow in a network” (Borgatti et al., 2009).

While getting these interpretations right is not trivial even within flow networks, it is a further problem when measures designed for analyzing other network mechanisms are applied. For example, using a grouping algorithm that has its logic in a similarity mechanism to data based on a logic of flow will lead to invalid conclusions. Mis-match between logics and algorithms means that “we lose the ability to interpret the measure…or we get poor answers” (p. 56). Getting such matching correct means grappling with the inter-connections between all the decisions we consider above.

Therefore, we recommend that researchers explicitly describe the mechanism they expect to see and use these as the basis for arguing for the overall cohesion of their network analysis decisions, arguing from theory at each decision. Researchers may find Borgatti’s taxonomy useful, or seek other authors who have concentrated on the links between networks and theoretically derived processes, such as Monge and Contractor (2003). Reviewers and editors may find referring authors to these contributions will assist the authors in making explicit their assumptions about network processes and the extent to which their network operationalizations validly capture these processes, providing the theoretical binding that joins the links in the chain of reasoning.

4. Conclusions

The combination of exciting phenomena based on digital interactions, copious data, interesting research questions, and appropriate methods, creates excellent opportunities for research. Social network analysis with digital trace data constitutes a “measurement revolution” (Kleinberg, 2008) because it provides a way of harnessing the data contained in online archives and using it to operationalize concepts of deep theoretical interest.

Nonetheless, this paper sounds a strong note of caution about the manner in which SNA concepts are translated to research using digital trace data. Through an analysis based on a detailed consideration of the types of data available and widely used, we have argued that digital trace data are of a different nature than those used in earlier studies using SNA. While there exists a literature on validity issues arising from these earlier methods, despite the surge in research using SNA with digital trace data, a corresponding validity literature has not emerged. This paper is a contribution to such a literature. It raises a set of pernicious validity concerns that extend throughout the links in the chain of reasoning, and thus, the decisions researchers must make to conduct network analysis, iterating from theoretical interests, to data collection, through initial transformation and reduction to networks, and following the chain of logic from construct, operationalization, and analysis of those networks. Information Systems researchers specifically have an excellent opportunity to contribute, drawing on their understanding of the contingent impact of systems, their grasp of structurational theories, and their particular interest in the phenomena generating these digital trace data.

By providing recommendations and highlighting studies that deal well with these challenges, we hope to improve the quality of SNA-based research using digital trace data, especially in terms of theoretical cohesion, and so position the field to make important contributions to the "twenty-first century science" of network analysis of online activity (Watts, 2007).

Acknowledgements

This research was partially supported by US National Science Foundation Human and Social Dynamics Program Grant 05-27457.

The first author would like to thank Carnegie Mellon University, especially James D. Herbsleb, for hosting him as a post-doctoral during revisions on this paper. The paper benefited from discussions with doctoral students in the CMU Computation and Society Ph.D. program, including George Davis, Jamie Olson, Peter Landwehr, Jana Diesner, Laurie Jones and Jesse St. Charles.

Thanks also to Sean Goggins for discussion and feedback on drafts of this paper.

References

Adamic, L., & Adar, E. (2005). How to search a social network. *Social Networks*, (27), 203.

Agarwal, R., Gupta, A. K., & Kraut, R. (2008). Editorial overview—The interplay between digital and social networks. *Information Systems Research*, *19*(3), 252.

Ahuja, M. K., & Carley, K. M. (1999). Network structure in virtual organizations. *Organization Science*, *10*(6), 757.

Aral, S., Brynjolfsson, E., & van Alstyne, M. (2006). Information, technology and information worker productivity: task level evidence. *International Conference on Information Systems (ICIS 2006)*. Retrieved from http://aisel.aisnet.org/icis2006/21

Bampo, M., Ewing, M. T., Mather, D. R., Stewart, D., & Wallace, M. (2008). The effects of the social structure of digital networks on viral marketing performance. *Information Systems Research*, *19*(3), 290.

Bird, C., Gourley, A., Devanbu, P., Gertz, M., & Swaminathan, A. (2006). Mining email social networks. *Proceedings of the 2006 international workshop on Mining software repositories* (pp. 137-143).

Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, *27*(1), 71.

Borgatti, S. P., & Foster, P. C. (2003). The network paradigm in organizational research: A review and typology. *Journal of Management*, *29*(6), 991-1013.

Borgatti, S. P., Carley, K. M., & Krackhardt, D. (2006). On the robustness of centrality measures under conditions of imperfect data. *Social Networks*, *28*(2), 136.

Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). *UCINET 6.0*. Natick: Analytic Technologies.

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science*, *323*(5916), 892-895.

Braha, D., & Bar-Yam, Y. (2006). From centrality to temporary fame: Dynamic centrality in complex networks. *Complexity*, *12*(2), 59-63.

Brynjolfsson, E., Malone, T. W., Gurbaxani, V., & Kambil, A. (1994). Does information technology lead to smaller firms? *Management Science*, *40*(12), 1628-1644. doi:10.1287/mnsc.40.12.1628

Burt, R. S. (1992). *Structural holes: The social structure of competition*. Cambridge, MA: Harvard University Press.

Butts, C. T. (2008). Social Network Analysis with sna. *Journal of Statistical Software*, *24*(6), 1-51.

Cataldo, M., Herbsleb, J. D., & Carley, K. M. (2008). Socio-technical congruence: a framework for assessing the impact of technical and work dependencies on software development productivity. *Proceedings of the Second ACM-IEEE international symposium on Empirical software engineering and measurement (ESEM '08)*. Kaiserslautern, Germany. (pp 2-11) doi:10.1145/1414004.1414008

Christley, S., & Madey, G. (2007). Analysis of activity in the open source software development community. *Proceedings of the 40th Hawaii International Conference on System Sciences (HICSS 2007)*. doi: 10.1109/HICSS.2007.74

Committee on Network Science for Future Army Applications, National Research Council. (2005). *Network Science*. Washington, D.C.: The National Academies Press.

Concas, G., Lisci, M., Pinna, S., Porruvecchio, G., & Uras, S. (2008). Open Source Communities as Social Networks: An Analysis of Some Peculiar Characteristics. *Proceedings of the 19th Australian Conference on Software Engineering (ASWEC 2008)* (pp. 387-391).

Contractor, N. S., Whitbred, R., Fonti, F., Hyatt, A., Jones, P., & O'Keefe, B. (2000). Structuration and self-organizing networks. *Paper presented at the Winter Organizational Science Conference*. Keystone, CO.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Rand-McNally.

Crowston, K., & Howison, J. (2005). The social structure of Open Source Software development teams. *First Monday*, *10*(2). Retrieved from http://firstmonday.org/issues/issue10_2/crowston/index.html.

Daniel, S. L., & Diamant, E. I. (2008). Network Effects in OSS Development: The Impact of Users and Developers on Project Performance. *ICIS 2008 Proceedings* (p. Paper 122). Presented at the International Conference on Information Systems. Retrieved from http://aisel.aisnet.org/icis2008/122.

de Nooy, W., Mrvar, A., & Batagelj, V. (2005). *Exploratory social network analysis with Pajek*. Cambridge University Press.

Ebel, H., Mielsch, L.-I., & Bornholdt, S. (2002). Scale-free topology of e-mail networks. *Physical Review E*, *66*. doi:10.1103/PhysRevE.66.035103

Falkowski, T., Barth, A., & and Spiliopoulou, M. (2008). Studying Community Dynamics with an Incremental Graph Mining Algorithm. *AMCIS 2008 Proceedings*. Retrieved from http://aisel.aisnet.org/amcis2008/29

Franz, R. S. (1998). Task interdependence and personal power in teams. *Small Group Research*, *29*(2), 226-253. doi:10.1177/1046496498292005

Freeman, L.C., Roeder, D., & Mulholland, R. R. (1979). Centrality in social networks: II. experimental results. *Social Networks*, *2*(2), 119-141.

Gao, Y., Van Antwerp, M., Christley, S., & Madey, G. (2007). A research collaboratory for open source software research. *Proceeedings of the First International Workshop on FLOSS (FLOSS '07)* (p. 4). doi:10.1109/FLOSS.2007.1

Garton, L., Haythornthwaite, C., & Wellman, B. (1997). Studying online social networks. *Journal of Computer-Mediated Communication*, *3*(1). Retrieved from http://jcmc.indiana.edu/vol3/issue1/garton.html

Geiger, R. S., & Ribes, D. (2011). Trace Ethnography: Following Coordination through Documentary Practices. *Proceedings of the 44th Hawaii International Conference on System Sciences (HICSS 2011)* (pp. 1 - 10). doi:10.1109/HICSS.2011.455

Giddens, A. (1984). *The constitution of society: outline of the theory of structuration*.

Goggins, S., Galyen, K., & Laffey, J. (2010). Network analysis of trace data for the support of group work: activity patterns in a completely online course. *Proceedings of the 16th ACM international conference on Supporting group work (GROUP '10)*. Sanibel Island, FL. doi:10.1145/1880071.1880089

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1380.

Grewal, R., Lilien, G. L., & Mallapragada, G. (2006). Location, location, location: How network embeddedness affects project success in open source systems. *Management Science*, *52*(7), 1056.

Grippa, F., Zilli, A., Laubacher, R., & Gloor, P. (2006). E-mail may not reflect the social network. *Proceedings of the North American Association for Computational Social and Organizational Science Conference (NAACSOS '06)*.

Hahn, J., Moon, J. Y., & Zhang, C. (2008). Emergence of New Project Teams from Open Source Software Developer Networks: Impact of Prior Collaboration Ties. *Information Systems Research*, *19*(3), 369-391. doi:10.1287/isre.1080.0192

Hansen, D., Shneiderman, B., & Smith, M. A. (2010). *Analyzing Social Media Networks with NodeXL: Insights from a Connected World*. Boston, MA: Morgan Kaufmann.

Hinz, O., & Spann, M. (2008). The Impact of Information Diffusion on Bidding Behavior in Secret Reserve Price Auctions. *Information Systems Research*, *19*(3), 351-368. doi:10.1287/isre.1080.0190

Howison, J. (2009). *Alone Together: A socio-technical theory of motivation, coordination and collaboration technologies in organizing for free and open source software development* (Ph.D. Dissertation). Syracuse University. Retrieved from http://james.howison.name/pubs/dissertation.html

Howison, J., Inoue, K., & Crowston, K. (2006). Social dynamics of free and open source team communications. *Proceedings of the International Conference on Open Source Software (IFIP OSS)* (pp. 319-330).

Huisman, M., & Snijders, T. A. B. (2003). Statistical analysis of longitudinal network data with changing composition. *Sociological Methods & Research*, *32*(2), 253-287.

Hume, D. (2000). *A Treatise of Human Nature*. Oxford University Press, USA.

Kane, G. C. (2009). It's a network, not an encyclopedia: a social network perspective on Wikipedia collaboration. *Best Paper proceedings of the Academy of Management annual meeting*. Chicago, IL.

Kane, G. C., & Alavi, M. (2008). Casting the Net: A Multimodal Network Perspective on User-System Interactions. *Information Systems Research*, *19*(3), 272.

Kazienko, P., Musial, K., & Kajdanowicz, T. (2008). Profile of the Social Network in Photo Sharing Systems. *AMCIS 2008 Proceedings*. Retrieved from http://aisel.aisnet.org/amcis2008/173

Kilduff, M., & Tsai, W. (2003). *Social networks and organizations*. Thousand Oaks, CA: Sage Publications Ltd.

Kilduff, M., Crossland, C., Tsai, W., & Krackhardt, D. (2008). Organizational network perceptions versus reality: a small world after all? *Organizational Behavior and Human Decision Processes*, *107*(1), 28.

Kleinberg, J. (2008). The convergence of social and technological networks. *Communications of the ACM*, *51*(11), 66-72.

Kossinets, G., & Watts, D. J. (2006). Empirical analysis of an evolving social network. *Science*, *311*(5757), 88-90.

Krackhardt, D., & Carley, K. M. (1998). PCANS model of structure in organizations. *Proceedings of the 1998 International Symposium on Command and Control Research and Technology*. Monterey, CA.

Lakhani, K., & von Hippel, E. (2003). How open source software works: "free"' user-to-user assistance. *Research Policy*, *32*(6), 923-943.

Latapy, M., & Magnien, C. (2008). Complex network measurements: Estimating the relevance of observed properties. *Proceedings of the 27th Conference on Computer Communications (IEEE INFOCOM 2008)* (pp. 1660-1668). Phoenix, AZ.

Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press.

Leskovec, J., Kleinberg, J., & Faloutsos, C. (2005). Graphs over time: Densification laws, shrinking diameters and possible explanations. *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining (KDD '05)* (pp. 177-187). New York, NY.

Long, Y., & Siau, K. (2007). Social network structures in open source software development teams. *Journal of Database Management*, *18*(2), 34-40. doi:10.4018/jdm.2007040102

Marsden, P. V. (1990). Network data and measurement. *Annual Review of Sociology*, *16*(1), 435-463. doi:10.1146/annurev.so.16.080190.002251

Merlo, F., Slaughter, S., & Francalanci, C. (2009). The Co-Evolution of Social Networks And Software Structures In Open- And Closed-Source Projects. *Best Paper Proceedings of Academy of Management Annual Meeting 2009*.

Monge, P. R., & Contractor, N. S. (2003). *Theories of Communication Networks*. Oxford: Oxford University Press.

Moody, J., McFarland, D., & Bender-deMoll, S. (2005). Dynamic network visualization. *American Journal of Sociology*, *110*(4), 1206-1241.

Nia, R., Bird, C., Devanbu, P., & Filkov, V. (2010). Validity of Network Analyses in Open Source Projects. *Proceedings of the Seventh Working Conference on Mining Software Repositories*.

Olson, J., & Carley, K. (2011). Reliability and Robustness of Network Analysis from Dynamic Social Events. *Proceedings from Sunbelt XXXI*. Trade Winds Beach Resort.

Opsahl, T., & Panzarasa, P. (2009). Clustering in weighted networks. *Social Networks*, *31*(2), 163.

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, *32*(3), 251.

Orlikowski, W. J. (1996). Improvising organizational transformation over time: A situated change perspective. *Information Systems Research*, *7*(1), 92.

Orlikowski, W. J., & Iacono, C. S. (2001). Research Commentary: Desperately Seeking the "IT" in IT Research: A call to theorizing the IT Artifact. *Information Systems Research*, *12*(2), 145.

Poole, M. S., & DeSanctis, G. (1990). Understanding the use of group decision support systems: The theory of adaptive structuration. In C. Steinfield & J. Fulk (Eds.), *Organizations and communication technology* (pp. 173-193). Beverly Hills, CA: Sage.

Rice, R. E. (1990). Computer-mediated communication system network data: Theoretical concerns and empirical examples. *International Journal of Man-Machine Studies*, *32*(2), 647.

Ridings, C. M., Gefen, D., & Arinze, B. (2002). Some antecedents and effects of trust in virtual communities. *Journal of Strategic Information Systems*, *11*(3-4), 271-295. doi:10.1016/S0963-8687(02)00021-5

Robey, D., Vaverek, A., & Saunders, C. S. (1989). Social Structure and Electronic Communication: A Study of Computer Conferencing. *Paper presented at the Hawaii International Conference on Information Systems, Hawaii.*
Sechrest, L. (2005). Validity of Measures Is No Simple Matter. *Health Services Research*, *40*(5p2), 1604.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Belmont, CA: Wadsworth Publishing.
Trier, M. (2008). Research Note---Towards Dynamic Visualization for Understanding Evolution of Digital Communication Networks. *Information Systems Research*, *19*(3), 350.
Valverde, S., Theraulaz, G., Gautrais, J., Fourcassie, V., & Sole, R. V. (2006). Self-organization patterns in wasp and open source communities. *IEEE Intelligent Systems*, *21*(2), 36-40.
Wagstrom, P. A., Herbsleb, J. D., & Carley, K. M. (2005). A Social network approach to free/open source software simulation. *Proceedings of the First International Conference on Open Source Systems (IFIP 2.13)*.
Wasko, M., & Faraj, S. (2005). Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice. *MIS Quarterly*, *29*(1), 57.
Wasserman, S., & Faust, K. (1994). *Social Network Analysis*. Cambridge: Cambridge University Press.
Watts, D. J. (2007). A twenty-first century science. *Nature*, *445*, 489. doi:10.1038/445489a
Wiggins, A., Howison, J., & Crowston, K. (2008). Social dynamics of FLOSS team communication across channels. *Proceedings of the Fourth International Conference on Open Source Software (IFIP 2.13)*.
Wu, J., & Tang, Q. (2007). Analysis of Survival of Open Source Projects: a Social Network Perspective. *PACIS 2007 Proceedings*. Retrieved from http://aisel.aisnet.org/pacis2007/19
Wu, J., Goh, K.-Y., & Tang, Q. (2007). Investigating Success of Open Source Software Projects: A Social Network Perspective. *ICIS 2007. Proceedings of International Conference on Information Systems 2007*.
Yeow, A., Johnson, S., & Faraj, S. (2006). Lurking: Legitimate or Illegitimate Peripheral Participation? *ICIS 2006 Proceedings*. Retrieved from http://aisel.aisnet.org/icis2006/62
Zwijze-Koning, K. H., & de Jong, M. D. T. (2005). Auditing Information Structures in Organizations: A Review of Data Collection Techniques for Network Analysis. *Organizational Research Methods*, *8*(4), 429-453. doi:10.1177/1094428105280120

About the Authors

James HOWISON is an Assistant Professor in the School of Information at the University of Texas at Austin. He holds a Ph.D. in Information Science and Technology from the Syracuse University School of Information Studies (2009). His research focuses on the organization of work on information technologies, particularly the organization of software development. He has pursued this in NSF-funded research on free and open source software development and the socio-technical organization of software production in science.

Andrea WIGGINS is a PhD candidate in Information Science and Technology at the School of Information Studies at Syracuse University. She received her MS in Information (2007) from the University of Michigan School of Information. Her research interests center on large-scale collaboration, and the focus of her dissertation is on the role of technologies in supporting public participation in scientific research.

Kevin CROWSTON is a Professor in the School of Information Studies at Syracuse University. He received his Ph.D. (1991) in Information Technologies from the Sloan School of Management, Massachusetts Institute of Technology (MIT). His research examines new ways of working and organizing made possible by the extensive use of information and communications technology. Specific research topics include the development practices of Free/Libre Open Source Software teams, work practices, and technology support for citizen science research projects and computational qualitative data analysis, all with NSF support.