Digital Rhetoric: Theory, Method, Practice
Douglas Eyman
Series: Digital Humanities

DOI: http://dx.doi.org/10.3998/dh.13030181.0001.001 [http://dx.doi.org/10.3998/dh.13030181.0001.001]

// three // Digital Rhetoric: Method

In addition to addressing the roles and activities of the speaker/writer, communication/text, and audience/reader, definitions of rhetoric that address digital communication need to account for context, interactivity, and circulation (via internetworked systems). Lloyd Bitzer's (1968) articulation of rhetoric as "a mode of altering reality, not by the direct application of energy to objects, but by the creation of discourse which changes reality through the mediation of thought and action" (4) provides a useful starting point for digital rhetoric by virtue of being an abstraction that does not explicitly address or evoke specific practices or media associated with rhetorical production while simultaneously acknowledging the power of rhetoric as a meaning-making activity. Hauser (1986) provides a more streamlined general definition of modern rhetoric as "the management of symbols in order to coordinate social action" (3); for both Bitzer and Hauser, rhetoric is an activity and not just an analytic framework.

In the case of the definitions I draw on above, none of the theorists address the complications of digital circulation or the possibilities of nonhuman agents becoming rhetorical actors. And while Hauser acknowledges that symbolic modes need not be constrained to the verbal, he does not address these other modes in his work (moreover, it is important to understand how multimedia and multimodality function at the intersection of multiple symbolic modes, and how this might complicate the "management of symbols"). Digital rhetoric, then, should take into account the complications of the affordances of digital practices, including circulation, interaction, and the engagement of multiple symbol systems within rhetorical objects, and its methods need to explicitly engage those complications and affordances.

Because I situate my professional identity at the same nexus as the point of origin for digital rhetoric as a field of study, I see composition/rhetoric, computers and writing, and professional writing as the fields that best understand how to research rhetoric and writing, and by extension, that provide the most effective starting points for assembling digital rhetoric methods. In this chapter, I will first address the traditional rhetorical method of close reading and the relatively new inverse of that method, which Franco Moretti (2000) calls "distant reading." I then cover the methods from fields in writing studies and then go on to examine methods from fields that do not take rhetoric as their theoretical or methodological foundations.

Close and Distant Reading as Rhetorical Methods

One of the most widely employed rhetorical methods is close reading. Close reading as a technique promoted by I. A. Richards (1930) focuses upon meaning within the text as it is evidenced in formal qualities (such as

rhythm, use of imagery and metaphor) as interpreted by the reader; in this version, the text is considered apart from the author, its cultural or historical context, or the material conditions of its construction; as Edwin Black (1965) notes, the aim is to determine "the purpose of a text from evidence the text itself provides" (16). Taken up as a foundational methodology by the New Critics in literary study, it has since shifted from methodology to method (i.e., technique) and the term has taken on the broader meaning of attentive reading in the sense that its formal qualities are reflections of social and historical effects and that the text itself may deploy rhetorical power outside of its internal interpretation. Barry Brummett (2010) defines close reading as "the mindful, disciplined reading of an object with a view to deeper understanding of its meanings" (3). It is in this sense that close reading, which Brummett connects to paying attention to the implicit contributions of the text to sociocultural effects (such as privileging a particular language over others) that is of the most use to digital rhetoric. The practice of engaging the formal qualities of a text can also be useful, but in this case "text" must be read in the comprehensive sense outlined in chapter 2 [http://quod.lib.umich.edu/d/dh/13030181.0001.001/--digital-rhetoric-theory-method-practice?g=dculture;trgt=div1_ch2;view=fulltext;xc=1] and the formal qualities would include those specific to different media (which may be disaggregated within the process of close reading for critical interpretation).

It may seem obvious that close reading or textual criticism is available as a method, but it has such strong ties to print text that I want to be sure that it does not lose ground as a method for digital text, particularly given the lengthy arguments that situate digital text as distinct from print text and my subsequent arguments that we need to develop digital-native methods for born-digital texts. Close reading, in the sense of applying our individual faculties to the interpretations of any given text, will nearly always be in play as an undercurrent of other methods.

But close reading also serves as a starting point because it provides the contrast for newer rhetorical methods; one of these, distant reading, can be seen as a natural opposite in terms of technique—rather than bounding the text and looking only at what it offers, distant reading takes a long view, examining the text as one among many and considering a much larger corpus whose contexts and relationships give rise to different forms of meaning. Franco Moretti's (2000) practice of "distant reading" sees distance as "*a condition of knowledge*: it allows you to focus on units that are much smaller or much larger than the text: devices, themes, tropes—or genres and systems" (57, emphasis in original). Distant reading methods require computational processes whose output is presented as specific forms of data visualization to dramatically alter the scales at which readers encounter texts (Mueller, 2009). Moretti (2005) examines the employment of three types of data visualization applied to large-scale corpus of literary texts: graphs, maps, and trees, and Derek Mueller (2009) adds clouds as an additional distant reading visualization (more on tag clouds appears later in this chapter, when I consider data visualization as method). One of the focal points of this chapter is to encourage synthesis and development of new methods for digital rhetoric, just as Moretti combined computational analytics and data visualizations in order to develop a new method for literary study—but also to encourage the extension of extant methods, as Mueller has done with Moretti's work.

Research Methods in Writing Studies

New texts on research methods in writing studies (including Bazerman [2007] and Hughes & Hayhoe [2007]) for the most part don't include digital methods. Individual scholars have been pushing for new methods that take into account the role of digital production in rhetorically centered research methodologies (see, for instance, Spinuzzi [2003], Spinuzzi & Zachary [2000], and Hart-Davidson [2005], discussed in more detail below,

as well as Swarts [2008] and Potts [2009]), but fewer researchers in composition and technical communication are focusing on tools and methodologies that arise from the rhetorical functions of the digital network itself.

Composition as a discipline is currently undergoing a significant shift in its overall focus: as composition continues to engage multiple modes and media as acceptable forms of composition (beyond the tradition of print-based writing), the practices and processes of composing that composition takes as its object of inquiry are undergoing radical changes—changes that necessitate concomitant changes in research methods. These changes amount to what is essentially an epistemological shift from a view of the solitary writer who has available only limited material means of production and often no recourse to distribution or circulation of the work, to a view of composition as a collaborative activity that engages multiple means of production and that occurs within digital networks that provide broad opportunities for publication and circulation.

The research methods in professional writing and technical communication tend to lend themselves more readily to the discovery of agents interacting in writing ecologies; Laura Gurak and Mary Lay's *Research in Technical Communication* (2002) contends that the foundational research methods in professional writing are "ethnography, textual analysis, historical research, survey and questionnaire research, and experimental work" (vii). The methods of professional writing, like composition/rhetoric and computers and writing, tend toward the qualitative, although the field is more accepting of quantitative methods and experimentation. Historically, professional writing research has paid more attention to context (particularly in terms of organizations and workplaces) than other writing studies research traditions.

Two of the key research traditions from professional writing that are particularly appropriate for digital rhetoric are genre studies and usability.

Genre studies, as elaborated in professional writing research, focus on investigations of "an individual's repertoire of situationally appropriate responses to recurrent situations" through examinations of the "situated actions of writers and the communication systems in which those . . . actors participate" (Berkenkotter & Huckin 1995, ix). In methodological terms, genre studies privileges a multilayered approach that engages both micro- and macro-level interactions. As Berkenkotter and Huckin explain,

> . . . what microlevel studies of actors' situated actions frequently depict as individual processes, can also be interpreted (from the macrolevel) as communicative acts within a discursive network or system. Genre is the concept that enables us to envision the interpretation of process and system in disciplinary communication. (ix–x)

This approach to the study of writing processes and practices is particularly useful when applied to digital environments, which engage individual and collaborative practices that take place within both digital and discursive networks. Focusing the lens on the activity of the writer or the context (and its conditional affordances for composing) allows a view that collapses system-centric and user-centric activity.

Another methodology that is especially well-suited to the study of digital composition is usability. Usability is not well understood as a rhetorically based qualitative research methodology outside of the field of professional and technical communication; more often than not, it is equated with observing users performing tests of preset activities under controlled conditions and is typically seen as *developmental* (i.e., typically developing

information tools, interfaces, and systems), and not as a research methodology at all. However, if usability is rearticulated as a method of investigating actual use in specific contexts and cultures, it is clear that it can be a powerful method for understanding rhetorical knowledge-making activity within a broad range of contexts and uses. As I've written elsewhere (Eyman, 2009),

> To engage usability as a suitable methodology for studying writing processes and pedagogies, it's important to first acknowledge that writing is a technology, and, consequently, that teaching writing is part of a technological system; a system with which our students interact as users. Constructing students as users allows us to see them not as subordinate to the learning process, but as engaged participants in the technological system that is bounded by the institutions, departments, and physical spaces in which learning activities take place. Students have particular needs and goals, but we don't always have a clear understanding of what those needs and goals are from the perspective of the user; curricular design is all too often enacted through a systems-design framework, rather than a user-centered framework. (222)

Usability, in other words, provides a methodology for studying both writing practices and writing pedagogies—and because it takes both system and user into consideration, it provides appropriate methods for studying digital writing practices and digital pedagogies.

Digital Writing Research

While many traditional research methods in composition/rhetoric and professional writing—particularly qualitative research practices—will

continue to function well regardless of the material conditions of production, new methods need to be developed to help us better understand how composing practices change from traditional print production activities to multimodal, multimedia productions that can now be delivered, distributed, published, and circulated in and through digital networks.

The general trend of research in composition/rhetoric and professional writing toward qualitative methodologies works well for the study of digital compositions because it takes into account situation, context, and media. Case studies, textual inquiry, and rhetorical analysis are particularly useful for investigations of rhetorical activity in digital environments, although in each case there is room for enhanced methods that can be adapted for use in digital networks. While the methods currently available cover quite a bit of ground in terms of researching digital writing practices, there are a few areas for which appropriate methodologies have not yet been developed, as well as a series of emerging methods that show a great deal of promise.

New methods include systems of visualizing discrete elements in the writing process as it takes place between and among multiple composers/authors. Hart-Davidson, Carter, and Sun (2006) suggest that producing different views (visual representations) of particular compositional and communicative activities can support different frames of analysis. This methodology is tied to a revision of the nature of composition as a rhetorical practice, as they assume "that writing is a medium, and that people are more often users of texts (as opposed to participants in a conversation); writing is not the focus of the action, but a powerful context for action" (20).

Shifting the research paradigm from a study of writing-as-action to writing-as-context allows for the development of new methods that might help us better see how this approach to the use of writing may be investigated.

William Hart-Davidson's (2005) work on establishing a rhetoric of objects, relationships, and views is an example of how context, system, and user might work well as the focus of inquiry for writing-as-context.

In the past decade, there have been few works that address digital, networked writing in terms of research methods—many articles and book chapters have explored the way that literacy changes when it takes place in digital contexts, how teaching must change to be successful for online courses, or have provided examples of new media practice—but the general consensus seems to be that we can apply traditional rhetorical, genre, or discourse analysis methods regardless of medium or context. In 2007, however, editors Heidi McKee and Dànielle DeVoss published the first collection that explicitly addresses methods for digital writing research: *Digital Writing Research: Technologies, Methodologies, and Ethical Issues*.

McKee and DeVoss (2007) define "digital writing research" as

> research that focuses: (a) on computer-generated, computer-based, and/or computer-delivered documents; (b) on computer-based text-production practices (and we deploy *text* broadly here, to include multimedia artifacts); and/or (c) on the interactions of people who use digital technologies to communicate. . . . Further, the term digital writing research—rather than the more commonly used term *Internet research*—acknowledges that not all digital writing and related communicative acts and interactions occur on the Internet. (3)

Digital Writing Research is an important collection for a number of reasons: it represents an acknowledgment of the broadening scope of what counts as writing activities—as James Porter notes in the foreword, "the

chapters in *Digital Writing Research* show us, either implicitly or explicitly, that the definition of "writing" has changed in the digital age and that, consequently, our approaches to doing research need to change; we need a parallel and equally dramatic change in our notions of methodology" (xiii). But the collection also represents a starting point for a disciplinary engagement with digital research methods for writing studies that has the potential to bring into sharp relief the kinds of theoretical and methodological shifts that must happen when writing moves from print to digital in nature. As Porter argues in the foreword:

> Likewise, digital writing research should not be viewed merely as research about writing with technology. It should be viewed, rather, as changing the fundamental assumptions about methodology, particularly the humanist assumption that divides the human from the technological. Digital writing research takes a cyborgian view and a networked view of human communications. It is not simply old methods applied to new events or practices. It represents a new way of looking altogether—an approach that emphasizes the role of production, delivery, and technology in human communication, but even beyond that, acknowledges the hybrid, symbiotic relationship between humans and machines. (xv–xvi)

While many of the approaches that appear in *Digital Writing Research* contribute to a rhetorical reconfiguration of the specific methods I examine below (and will thus be addressed within the contexts of those methods), the works by Kevin DePew and Julia Romberger in *Digital Writing Research* together provide a framework within which all of the following methods may be employed.

DePew (2007) argues for the importance of triangulation—of looking not just at texts but at contexts and users. He suggests that, "as rhetoricians, we should be examining more features of the communicative situation rather than merely an artifact it produces. What else can we learn about digital rhetoric when we also study the rhetor's intentions? The audiences' response to the text? How local contexts shape this interaction? . . . In essence, I am advocating that digital rhetoric researchers adopt strategies framed by the communicative triangle—the rhetor, the audience, the digital text or discourse, and the contexts. By designing such methodological strategies, researchers insert communicative participants into the process, which gives researchers the opportunity to see both the complex nature of the research site and apertures in the field's tropes" (52). I would add that the communicative participants need not be solely human audiences, but may be elements of the networked digital systems themselves (indeed, Jason Swarts [2008] provides an example of how nonhuman discursive agents may be included in technical writing research methods).

Romberger's (2007) work similarly focuses on context but addresses it within an ecological metaphor:

> An ecofeminist methodology, in short, must be aware of context and its complexity—the ecology of the situation. It is this emphasis on the influence of environment upon subjects in an ecological ebb and flow and how these relationships are articulated that separates it from other feminist methodologies. It takes into account histories of the larger social milieu and remains aware of the context of the researcher and the system of values brought in by framing an inquiry in a specific theory and discipline" (250).

Taken together, these two positions—engaging context and expanding the

scope of research methods to include the textual, the social, and the rhetorical situation—provide a platform for digital rhetoric research. But before such a platform can be fully articulated and deployed, the methods that work within it must be identified and, in some cases, developed. In the remainder of this chapter I review research methods from a broad range of fields and disciplines that may be profitably appropriated for digital rhetoric research.

C.O.D.E. and Network Administration Tools

If rhetoricians are to develop methods that are "digital-native," then looking to the tools and metrics that run these digital environments, such as network and routing tools and the protocols upon which the Internet was built, would be a logical first step. Even though digital texts are themselves immaterial, the networks in which they reside are made of physical data conduits and routing devices. When I worked as a webmaster and systems administrator for a community college, I learned about a variety of tools that were designed to monitor the health and productivity of these networks: I could keep track of how many hackers were attempting to infiltrate my servers or how robust the network link between two buildings was on a moment-by-moment basis. Although I know of several technorhetoricians whose backgrounds include experience in systems administration or programming, the first to articulate a coherent method for using these network tools for digital writing research is James Ridolfo (2006), who developed a webtext evaluation suite that he called "C.O.D.E.—Comprehensive Online Document Evaluation." Ridolfo presents this suite of tools as a pedagogical application that students can use to "not only cite online documents, but also critically research . . . digital texts."

Ridolfo provides instruction on using three network analysis tools to uncover both geographies and owners of digitally networked systems, along

with two additional web-based tools for examining the changes over time that a given website experiences. The "geography and ownership" tools that Ridolfo discusses include whois, traceroute, and ARIN. When I first learned to use tools like traceroute, the only way to do so was via the command line (usually on another server, although these tools are available on all personal computers as well). However, web-based interfaces for these tools have been developed—making the tools themselves more accessible to students and researchers alike. The first tool in C.O.D.E. is called "whois." Whois (http://www.betterwhois.com [http://www.betterwhois.com]) allows the user to retrieve information about who has registered a domain name, including date of registration, administrative contacts, and billing addresses. Traceroute, the second tool in C.O.D.E., traces the route that an Internet request must travel to reach its destination. For instance, when you use a web browser to view a page such as http://www.msu.edu [http://www.msu.edu] , your browser sends a request from wherever you are to the server that hosts that site; this request travels through the various hubs and routers that lay between your computer and the server at Michigan State University. Traceroute reinforces the geographic nature of interconnected networks and generally shows the overall distance between two networked points. The output of traceroute also shows the names of the routers and systems it encounters, so you can learn which Internet Service Provider (ISP) hosts the server at the end of the trace. Ridolfo argues that "these two utilities allow us to . . . contextualize the website based on its geographic origins, 'publisher' (ISP), time, and new authorial information" (n.p.). The final tool in the C.O.D.E. suite is ARIN—American Registry for Internet Numbers (http://www.arin.net [http://www.arin.net])—which allows users to look up the registration information of Internet addresses. So, for instance, ARIN reports that the IP address 35.8.10.26 belongs to Michigan State University and that MSU's ISP is Merit Network Inc.

The other techniques covered in C.O.D.E. help the user to find out more

about the *web-based* (as opposed to the physical network-based) context of a given site. Using the Web Archive (http://archive.org [http://archive.org]) allows the researcher to access previous versions and edits of many websites. The original website for the online journal *Kairos: A Journal for Teachers of Writing in Webbed Environments* (which was renamed *Kairos: A Journal of Rhetoric, Technology, and Pedagogy* and changed web addresses in 2001) is available by searching for the original web address (http://english.ttu.edu/kairos [http://english.ttu.edu/kairos]) in the Web Archive; the archive also provides links to all of the versions and updates that have taken place since then. The final activity in C.O.D.E. is a Google search of the URL for the site under consideration. Searching for the URL (as opposed to the site name or content) provides a quick view of the role this site has in the larger discourse of online communities (however, it is not as powerful a tool as cataloging the links to that site from other sources, a technique that is discussed in the section below on bibliometrics and cybermetrics).

Because digital communication can be deceptively ethereal, these tools help to recover the underlying material structures of the digital networks we study; additionally, these methods also reveal the activities of the nonhuman actors in the system, such as the routers that carry and promote the network's communication signals and the servers that respond to the queries initiated by people or other servers.

Studying Web Usage via Server Log Analysis

A great deal of information about users of digital genres (such as blogs, wikis, or websites) can be found in the log files automatically generated by the servers that house digital texts. These logs record how a user's search strategy leads to a particular text, and how many individuals have accessed a given text. Server log analysis can show which pages are entry points for

users and which are exit points, how many times a given page is viewed, how many "unique users" have visited a site, and some basic information about where those users come from. It is possible to combine server log analysis with the use of cookies or content-management system supported sites to track how long users spend on a given page and what paths a user takes when moving through a site (server logs can also record what link or search engine result lead a user to a site's entry point, although it typically does not have access to the search query).

Server log analysis yields very basic quantitative data that can show how a specific site's traffic has changed over time, as well as some characteristics of a site's audience. Examining the server logs from the online journal *Kairos*, for instance, provides a picture of a steadily growing number of accesses over the past decade, as well as an increase in international audiences:

> In addition to the steady increase in overall readership, we've seen a shift from a primarily US audience to a much more international audience. A little over 80% of our readers come from the US, which means that about 20% come from elsewhere—the logs have recorded visitors who hail from 190 different country codes, from Belize, Belarus, Botswana and Brazil; from Vietnam, Venezuela, and the Ukraine. And that 20% is now over 9,000 readers—so I'd say it would be safe to consider *Kairos* an "international" scholarly publication venue. (Eyman, 2006)

Although server log analysis is limited, it can serve as a starting point for understanding the relationship between a given digital text (or context) and its audiences. Additionally, server log analysis provides data that can be used to help triangulate findings from other methods. Server log analysis is

tied to circulation analysis because it can provide a general picture of the number of individuals accessing a digital-native text and also provide some information about where those individuals are from. However, this kind of overview should be considered secondary information because it does not directly connect the digital texts to its users and uses. One significant drawback of server log analysis, however, is that the researcher needs to have access to the server logs themselves—and this kind of information resource is rarely made public.

Social Network Analysis (SNA)

Because of its focus on networked relationships and their support of the circulation of social capital, Social Network Analysis (SNA), a research approach from sociology and communication science, provides a powerful set of tools for digital writing research. Social network analysis focuses on patterns of relations among people, organizations, states—in other words, human relationships, but rarely human/nonhuman interactions or relationships (Wellman & Berkowitz, 1988; Scott, 1991; Wasserman & Faust, 1994). Social network analysis takes a mixed-method approach: SNA makes use of qualitative data gathered via interviews, surveys, observation, and artifacts (Rogers, 1987; Garton et al., 1997), but it uses quantitative analyses to interpret that data. As Lin Freeman (1997) notes,

> From the outset, the network approach to the study of behavior has involved two commitments: (1) it is guided by formal theory organized in mathematical terms, and (2) it is grounded in the systematic analysis of empirical data. It was not until the 1970s, therefore—when modern discrete combinatrics (particularly graph theory) experienced rapid development and relatively powerful computers became readily available—that the study of social networks really

began to take off as an interdisciplinary specialty. (n.p.)

The basic premise of social network analysis is that relationships cannot be discretely quantified as units of measurement; that is, the relationship between two individuals must always be seen within the context of all the other relationships those individuals engage in (either shared or separately). This approach presents a high level of complexity that is handled by statistical analysis and the mathematical formulas that describe networks in terms of nodes and ties; as Joseph Barnes (1972), credited with being the first researcher to study social networks, explains, "to discover how A, who is in touch with B and C, is affected by the relation between B and C . . . demands the use of the network concept" (3).

In social network analysis, nodes represent the individual actors within networks; ties represent the relationships shared by those actors—these relationships (also called "strands") can be described in terms of content (the resource that is exchanged), direction, and strength. Some network analysts have applied social network methods to electronic texts, using SNA tools to surface patterns of relations between words and phrases; however, unlike the kind of mapping that similar work in applied linguistics produces, SNA textual analysis is used to "reveal cognitive maps and identifies people who hold similar conceptual orientations" (Garton et al., 1997, n.p.).

Social networking analysis methods have been used to trace the circulation of social capital (Ooka & Wellman, 2003; Huysman & Wulf, 2004) and thus are particularly well-suited to questions of digital economies and circulation: as Barry Wellman (2003) notes, "Networks are a major source of social capital mobilizable in themselves and from their contents" (n.p.). The work that social analysts do focuses on tracking and tracing the movement of resources between people; they "seek to describe networks of

relations as fully as possible, tease out the prominent patterns in such networks, trace the flow of information (and other resources) through them, and discover what effects these relations and networks have on people and organizations" (Garton et al., 1997, n.p.).

Several researchers in rhetoric and writing have begun adapting social network analysis methods for studies of online interaction that are based on writing practices; these methods, however, are event-based rather than relationship-based (Hart-Davidson, 2007).

Hypertext Network Analysis (HNA)

Hypertext Network Analysis (HNA) is, in a sense, a form of social network analysis, but it moves the question of relationships away from people and organizations and instead looks at the nodes and ties of digital texts as instantiated in websites and web links. The key distinction between social network analysis and hypertext network analysis is that the websites themselves are considered actors within the networks being investigated: "In particular, through a hyperlink, an individual website plays the role of an actor who could influence other website's trust, prestige, authority, or credibility" (Park, 2003, 53).

Park and Thelwall (2003) argue that "compared to other Web methods such as a content-based analysis, the relative advantage of hyperlink analysis is that it is able to examine the way in which Web sites form a certain kind of relations with others via hyperlinks" (n.p.)—thus the hypertext link serves as the focal point of the investigation. Hypertext link analysis also tends to be applied to very large-scale data sets. Broder et al. (2000), for instance, examined two hundred million pages and 1.5 billion hyperlinks in a study that showed that the probability of a hyperlink path between two randomly chosen Web pages was about 24 percent. When a path was present, there

was an average of approximately sixteen hyperlinks in the path between pages. These kinds of topological investigations take advantage of network analyses in ways that are similar to those of Ridolfo's C.O.D.E. suite of networking tools, but they use only the explicit links among and between websites to uncover the connections between them.

Hyperlink analysis has also been applied as part of the methods available to cybermetrics, drawing on Rousseau's (1997) analogy between citations and hyperlinks (coining the term "sitation" to foreground the similarities). As Park and Thelwall (2003) note, "the analogy between hyperlinks and citations has continued to generate interest within information science, including speculations about the kind of information that they could reveal in different contexts" (n.p.). This connection of citation and hyperlink also evokes the circulation of social capital, as both hyperlinks and citations can be indicators of (and can be mapped as) social/academic capital forms of resource exchange.

Bibliometrics and Cybermetrics

The most obvious (and traditional) method of tracing the use and value of texts is through citation analysis, although its use is limited when considering the overall circulation of a text. Still, as Kaplan and Nelson (2000) point out, "in the absence of a more compelling metric, citation analysis remains the best commonly available indicator of usage" (324).

Citation analysis as a process and a field of study provides numerous means and methodologies for use in quantifying a record and history of citation for authors, articles, and journals. The simplest method of citation analysis is to select a time frame and a body of citation data and determine how many times an author, article, or journal has been cited by the publications indexed in the dataset within that time frame. In most cases, citation data

for this sort of bibliometric analysis is drawn from citation databases, such as Social Sciences Citation Index (SSCI), Science Citation Index (SCI), and Arts and Humanities Citation Index (AHCI), which are all accessible online from Thomson's Institute for Scientific Information (ISI) database, also known as the Web of Science.

Scholars also employ citation analysis methods called co-citation and author co-citation in order to map disciplines (Small, 1999; White & McCain, 1998), determine subfields within major areas of study (Bayer, Smart, & McLaughlin, 1990), and locate cross-disciplinary influences (Small, 1999). The raw data included in co-citation analyses of articles, journals, and authors includes the number of times that pairs of articles, journals, and/or authors are co-present in the works cited or footnotes of articles located through citation databases. As Bayer, Smart, and McLaughlin (1990) explain, co-citation assumes "that the more frequently two scientists are cited together, and the more similar their patterns of co-citations with others, the closer the relationship between them" (444). This kind of relationship can be viewed as an instance of circulation activity that can be directly tracked. The problem of utilizing bibliometric methods for examining circulation (both print and digital) is that the databases are not complete—they are selectively populated both in terms of the works and citations they track and by a calculation of value (in terms of academic capital) that is applied to those works based on the citations they receive from other works that already have an established value. Thus the scope is very limited in terms of an overall picture of knowledge production and circulation.

Of course, qualitative methods of citation analysis are also employed in order to determine how authors incorporate citations and the ideas of the texts they cite within their scholarship. Such analyses require examining the use of citations within scholarly texts to determine the rhetorical functions

of those citations (Budd, 1999; Hyland, 1999). Budd's (1999) study of internal citations in seventy sociology articles from 1990 to 1995 reveals that authors use most of their citations quantitatively, also called procedurally in Budd's terminology, in order to prove to readers that they thoroughly researched their respective topics and are aware of disciplinary contexts (271). As Budd notes, procedural citations, those not integral to knowledge claims made by the authors, outnumber epistemic citations by a ratio of more than three to one (271). Authors' use of citation in largely procedural ways supports our assertion above that the citation of particular materials reflects and reinforces the significance of those items as important texts in the field that must be taken into account and acknowledged by authors as a condition for the perceived credibility of their arguments, even if the references cited are not integral to their arguments. On the other hand, items that are not cited can be viewed as having less credibility and may be judged as largely irrelevant.

A more promising method for digital writing research can be located in new informetric methodologies—based in part upon the principles and statistical formulas developed for bibliometric analyses—that are being developed by researchers in the field of information science. Several terms for these new methodologies have been suggested, but the field currently appears to favor "cybermetrics" as the designation for the study of online scholarship.

Cybermetrics studies the network of links between electronic scholarly works, revealing how widely a specific electronic source is linked to other online texts, what types of texts link to specific sources, and how the source is used. Aguillo (2003) locates cybermetrics at the intersections of "cybergeography" and "cyberdemography" across Internet genres (such as e-mail, the World Wide Web, and online databases). Methods include adaptations of bibliometrics, user studies, calculations of "cyberindicators" (website hits, search engine rankings), assessment of web data architecture

and hyperlink topologies, and comparative search engine analyses.

Initial research on web linking began with bibliometric approaches, but it soon became apparent that new methodologies would need to be developed in order to study the web from an informetrics perspective: Larson (1996) used linking as an analogous method of citation analysis to devise a map of the intellectual structure of cyberspace; Kleinberg (1999) demonstrated that useful information about individual web pages and websites can be extracted directly from link structures; and Broder et al. (2000) asserted that hyperlinks themselves can be studied as objects of interest in their own right.

Content Analysis

Content analysis is the systematic, quantitative analysis of communication content (including verbal, visual, print, and electronic communication). According to C. W. Roberts in the *International Encyclopedia of the Social and Behavioral Sciences* (2001), "content analysis is a class of techniques for mapping symbolic data into a data matrix suitable for statistical analysis" (2697); in this regard, content analysis is similar to social network analysis, except that it focuses on the representations in and across individual texts rather than the relationships between them. List (2005) makes clear that content analysis, "though it often analyzes written words, is a quantitative method. The results of content analysis are numbers and percentages. After doing a content analysis, you might make a statement such as '27% of programs on Radio Lukole in April 2003 mentioned at least one aspect of peacebuilding, compared with only 3% of the programs in 2001'" (kya16a.html).

Content analysis is typically applied in one of two general modes: conceptual analysis or relational analysis. Conceptual analysis establishes

the existence and frequency of concepts—most often represented by words or phrases—in a text; in contrast, relational analysis examines the relationships among concepts in a text (Busch et al., 2005).

Busch et al.'s (2005) description of relational analysis echoes the call for understanding relation complexities that occur in social network analysis, where the individual ties have no meaningful relationship except within the context of the larger network:

> Relational analysis, like conceptual analysis, begins with the act of identifying concepts present in a given text or set of texts. However, relational analysis seeks to go beyond presence by exploring the relationships between the concepts identified. Individual concepts, in and of themselves, are viewed as having no inherent meaning. Rather, meaning is a product of the relationships among concepts in a text. (n.p.)

There are two forms of relational analysis that hold promise for digital writing research: proximity analysis and cognitive mapping. Proximity analysis, like co-citation analyses in bibliometrics, looks for the co-occurrence of concepts in the texts being studied. In text-based proximity analysis, the concept takes the form of a string of words. Cognitive mapping uses the results of a proximity analysis and displays them as a visual map that represents the relationships between concepts (this is, indeed, very similar to the sociograms of social network analysis, which provide maps of relationships between people or groups). Busch et al. (2005) enumerate the theoretical assumptions that support this kind of mapping: "mental models are representations of interrelated concepts that reflect conscious or subconscious perceptions of reality; language is the key to understanding these models; and these models can be represented as networks" (n.p.). These kinds of maps are difficult to create by hand; like the mathematical

approaches employed in social network analysis, the development of concept mapping for content analysis has been greatly aided by advances in computing—in other words, the digital environment itself is necessary to support these methods. Early proponents of concept mapping describe it as "a computerized multidimensional scaling technique that generates maps of content themes based on the frequency and co-occurrence of key words" (Miller & Riechert, 1994, 3).

One example of content analysis applied to digital writing research is Herring et al.'s (2004) "Women and Children Last: The Discursive Construction of Weblogs," in which the authors use content analysis techniques to assess the age and gender of weblog authors:

> Gender of blog authors was determined by names, graphical representations (if present), and the content of the blog entries (e.g., reference to "my husband" resulted in a "female" gender classification, assuming other indicators were consistent). Age of blog authors was determined by information explicitly provided by the authors (e.g., in profiles) or inferred from the content of the blog entries (e.g., reference to attending high school resulted in a "teen" age classification). (n.p.)

Herring et al. also used a content analysis rubric to develop type categories for the weblogs themselves, dividing them into "filters," which primarily feature links to world events, online happenings, and other nonauthor-centered issues; "personal journals," which primarily contain the blogger's thoughts and internal workings; and "k(nowledge)-logs," which are "repositories of information and observations with a typically technological focus" (n.p.). This second move is a rhetorically-informed variation on traditional content analysis techniques, which often do not take into account the context of the texts under consideration.

In our chapter in *Digital Writing Research*, Colleen Reilly and I utilized a similar form of content analysis to develop a heuristic for evaluating digital texts in terms of their structure, the digital environments in which they reside, and the degree to which they violate traditional print-based genre norms (Reilly & Eyman, 2007). Drawing on Bolter and Grusin's (1999) theories of remediation, we examined the content of electronic scholarly publications to determine their degree of departure from the conventions of print texts and the extent to which they exploit and even highlight the affordances, structure, and multimedia nature of texts native to digital environments. We established a continuum that includes four designations for electronic texts: highly transparent, moderately transparent, moderately hypermediated, and highly hypermediated (Reilly & Eyman, 2007). In order to code the websites that we examined as falling into these categories, it was necessary for us to consider not only textual content but also paratextual content (links, document structures) and the kinds of visual and interactive content that can be published on the web. Thus, our work is also an example of how content analysis techniques can be applied to both textual and visual elements in digital texts.

Data Visualization

Composition/rhetoric as a field is experiencing a renewed interest in the role of the visual, particularly as it is used in multimedia and multimodal compositions; professional and technical writing has long understood the importance of visual rhetoric for effective communication. The "turn to the visual" also plays a prominent role in digital research methods, particularly in the form of data visualization.

Visualization is not simply a tool for displaying the results of analytic methods; it is itself a method that can be used to structure data in ways that reveal patterns—in other words, it is an analytic technique in its own right.

Lengler and Eppler (2007) define visualization methods as "systematic, rule-based, external, permanent, and graphic representations that depict information in a way that is conducive to acquiring insights, developing an elaborate understanding, or communicating experiences" (n.p.). In their "Periodic Table of Visualization Methods for Management," Lengler and Eppler divide visualization methods into six distinct groups: data visualizations (visual representations of quantitative data in schematic form); information visualization (interactive visual representations of data designed to amplify cognition by transforming the data into an image that is mapped to screen space); concept visualization (2-D graphical displays where concepts are connected by directed arcs encoding brief relationships between pairs of concepts); metaphor visualization (which first positions information graphically to organize and structure it and then conveys an insight about the represented information through the key characteristics of the metaphor that is employed); strategy visualization (which uses complementary visual representations to improve the analysis, development, formulation, communication, and implementation of strategies in organizations); and compound visualization (a mix of any of the foregoing visualization types).

An example of data visualization used methodologically is Kichiyoshi et al.'s (1999) "Data visualization for supporting query-based data mining," which describes how visualization can help users test hypotheses about the structures and contents of databases with which they are interacting. In their method, "an instance in a database which has several attributes with numerical or nominal values is visualized as a color bar with several color parts which correspond to attribute values. Values of a function which evaluates the utility of a hypothesis are also visualized by using colors. This visualization technique helps users find an initial hypothesis and modify the hypothesis in order to increase the usefulness of it interactively" (888). Like this and other examples, most applications of data visualization as research

method come from quantitatively oriented disciplines, such as the use of sociograms in social network analysis and hypertext network maps in hypertext network analysis.

Data visualization is very useful for making accessible large-scale systems that might not otherwise be comprehensible. As Roger Brown (1965) explains,

> Social structure becomes actually visible in an anthill; the movements and contacts one sees are not random but patterned. We should also be able to see structure in the life of an American community if we had a sufficiently remote vantage point, a point from which persons would appear to be small moving dots. . . . We should see that these dots do not randomly approach one another, that some are usually together, some meet often, some never. . . . If one could get far enough away from it human life would become pure pattern. (165)

Applying the functionality of a concept map coupled with a frequency analysis of terms that appear in the journal *College Composition and Communication*, Derek Mueller (2007) has implemented a system that visually represents the main themes of each issue in the form of a "tag cloud."

Tag clouds can be generated automatically by extracting the most common phrases from a given corpus (as with the example above), or they can represent the tags that individuals apply to content in folksonomic systems. In either case, this form of data visualization has clear potential as a digital rhetoric method.

Complicating Factors for Digital Research Methods

Although each of the preceding methods can, I believe, be adapted, appropriated, and synthesized for use in digital rhetoric research (after infusing them with a rhetorical foundation and vision), there are several complicating factors that will affect any method used for digital writing research. The main factor (and the one from which the others derive) is access. Accessibility can be impeded by intellectual property gatekeeping (restricted access to networks and texts that circulate in and through those restricted systems, as well as cost-prohibitive access fees on certain content), but it is also an issue when considering the format of the rhetorical objects themselves. Digital texts can be made up of, in part, proprietary formats; they may also engage media or genres for which we currently have few (if any) tools that would allow us to fully understand how they work or their relationship to their digital environments. Finally, the problem of ephemerality is also endemic to digital texts: websites are not stable entities that are fixed (they are unfixed by their very nature), and many become inaccessible by ceasing to exist. As noted above, some sites are still recoverable via the Web Archive, but this is not always the case: Colleen Reilly has recently alerted me to the fact that the Web Archive now retroactively obeys no-index restrictions placed in robots.txt files, meaning that any site that adds that directive to a new version will cause all past versions to go dark in the archive. Quite a number of academic institutions and departments are applying these no-index controls (the rationale for doing so is unclear), which means that they are effectively erasing their histories from the network. Clearly, any digital-native methodology must be aware of how these issues of access and control constrain the methods that are used.