Web Sphere Analysis and Cybercultural Studies

Kirsten Foot

Department of Communication

University of Washington

One way of approaching cybercultural studies is to focus on the relations and patterns, means and artifacts of cultural production and exchange online. Viewed as an evolving set of structures that enable and manifest the production of cyberculture, the hyperlinked, co-produced and ephemeral nature of the Web challenges traditional approaches to research of social, political and cultural interchange. Cultural studies of the Web may benefit from new methods of analyzing Web form and content, along with processes and patterns of production, distribution, usage and interpretation of Web-based phenomena. In this chapter, I propose the concept of a *Web sphere* as a unit of analysis for cybercultural studies, explain the value of Web archives, and discuss methods of Web sphere analysis that may be useful for understanding cybercultural phenomena. I illustrate these methodological reflections through two studies of personal expression on the Web in the wake of the attacks of September 11, 2001.

Borrowing a concept from the work of Taylor and van Every (2000) on the relationship between communication and organizing, the Web could be viewed as both a "site and surface" for communicative action. In order to conduct developmental analyses of both the "site" and "surface" of the Web, it is helpful to create and analyze an archive not just of discrete Web sites but of a Web sphere. A Web sphere is a collection of dynamically defined digital resources spanning multiple Web sites deemed relevant or related to a central theme or object, in the sense

of the *gegenstand* concept from classical German philosophy (Foot and Schneider, 2002). The *gegenstand* notion of object as a focal point embedded-in-activity, (see Leont'ev, 1978, Foot, 2002), enables the identification of a Web sphere as a collaborative production. As a unit of analysis, a Web sphere as a unit of analysis is boundable by time and object-orientation, and is sensitive to developmental changes, within which social, political and cultural relations can be analyzed in a variety of ways.

The most crucial element in this definition of Web sphere is the dynamic nature of the sites to be included. This dynamism comes from two sources. First, the researchers involved in identifying the boundaries of the sphere are likely to continuously find new sites to be included within it. Second, as will be discussed below, the notion of defining a Web sphere is recursive, in that pages that are referenced by other included sites, as well as pages that reference included sites, are considered as part of the sphere under evaluation. Thus, as a Web sphere is archived and analyzed over time, it boundaries may be dynamically reestablished by both the researchers and the sites themselves. The Web sphere can function as a macro unit of analysis, by which historical and/or inter-sphere comparisons can be made. For example, the Web sphere of the 2000 elections in the United States can be comparatively analyzed with the U.S. electoral Web sphere of 2004, as well as with electoral Web spheres in other countries. Alternatively and/or simultaneously, other, more micro units, such as features, links or textual elements, can be employed in analyses of a Web sphere, as I explain below.

Web sphere analysis is an analytic strategy that, fully implemented, includes analysis of the relations between producers and users of Web materials, as potentiated and mediated by the

structural and feature elements of Web sites, hypertexts and the links between them (Schneider and Foot, 2005, Schneider and Foot, 2004). In a nutshell, the multi-method approach of Web sphere analysis consists of the following elements. Web sites related to the object or theme of the sphere are identified, captured in their hyperlinked context, and archived with some periodicity for contemporaneous and retrospective analyses. The archived sites are annotated with human and/or computer-generated "notes" of various kinds, which creates a set of metadata. These metadata correspond to the unit(s) and level(s) of analysis anticipated by the researcher(s). Sorting and retrieval of the integrated metadata and URL files is accomplished through several computer-assisted techniques. Interviews of various kinds are conducted with producers and users of the Web sites in the identified sphere, to be triangulated with Web media data in the interpretation of the sphere.

From the perspective of Web sphere analysis, the essence of the Web is the link (Foot et al., 2003). Links provide the nutrients that give the Web the energy and nourishment necessary for growth and development. Links serve as the neural pathways through which the collective intelligences and performances of Web producers and users are created, displayed and distributed. Several approaches have emerged that takes hyperlink relationality into account in more nuanced ways. Lindlof and Shatzer (1998) point in this direction in their article calling for new strategies of media ethnography in "virtual space." Hine (2000) presents a good example of sociocultural analysis of cross-site action on the Web. Similarly, Howard's (2002) conceptualization of network ethnography reflects methodological sensitivity to processes of Web production. In these examples and in Web sphere analysis, attention is given to the hyperlinked context(s) and situatedness of Web sites, and to the aims, strategies and identity-

construction processes of Web site producers, as they are produced, maintained and/or mediated through links.

In order to engage in any kind of developmental or retrospective study of cyberculture on the Web, it is helpful to capture Web materials in a time-sensitive way. The ongoing evolution of the Web poses challenges for scholars as they seek to develop methodological approaches permitting robust examination of Web phenomena. Some of these challenges stem from the nature of the Web, which is a unique mixture of the ephemeral and the permanent. There are two aspects to the ephemerality of Web content. First, Web content is ephemeral in its transience, as it can be expected to last for only a relatively brief time. From the perspective of the user or visitor (or researcher), specialized tools and techniques are required to ensure that content can be viewed again at a later time. Second, Web content is ephemeral in its construction -- like television, radio, theater and other "performance media" (Hecht et al., 1993, Stowkowski, 2002). Web content, once presented, needs to be reconstructed or re-presented in order for others to experience it. Some older media – including printed materials, film and sound recordings, for example -- can be archived in the form in which they are presented; no additional steps are needed to re-create the experience of the original. Although Web pages are routinely reconstructed by computers without human intervention (when a request is forwarded to a Web server), it nevertheless requires some action by the producer (or the producer's server) in order for the content to be viewed again. In other words, the experience of the Web, as well as the bits used to produce the content, must be intentionally preserved in order to be reproduced (Arms et al., 2001).

At the same time, the Web has a sense of permanence that clearly distinguishes it from performance media. Unlike theater, or live television or radio, Web content must exist in a permanent form in order to be transmitted. The Web shares this characteristic with other forms of media such as film, print, and sound recordings. The permanence of the Web, however, is somewhat fleeting. Unlike any other permanent medium, a Web site may regularly and procedurally destroy its predecessor each time it is updated by its producer. That is, absent specific arrangements to the contrary, each previous edition of a Web site may be erased as a new version is produced. By analogy, it would be as if each day's newspaper was printed on the same piece of paper, obliterating yesterday's news to produce today's.

The ephemerality of the Web requires that pro-active steps be taken in order to allow a recreation of Web experience for future analyses. The permanence of the Web makes this eminently possible. Although saving Web sites is not as easy as, say, saving editions of a magazine, archiving techniques are evolving in such a way to facilitate scholarly research of Web sites. In distinction to other ephemeral media, the Web can be preserved in nearly the same form as it was originally "performed," (Kahle, 1997, Lyman and Kahle, 1998, Lyman, 2002), and analyzed at a later time. Web archiving enables more rigorous and verifiable research, as well as developmental analyses that are time sensitive (e.g. Foot et al., 2003).

Robust Web sphere analysis benefits from robust Web archives. Going further, I suggest that Web archives enable an expanded range of investigable questions and greater analytical rigor for social research on Web-based phenomena. In the remainder of this essay I illustrate the potential affordances of Web archiving and Web sphere analysis for cybercultural studies through a brief

overview of the September 11 Web Archive project in general, and two studies of personal expression in the post-September 11 Web sphere that exemplify some of quantitative and qualitative methods of analysis enabled by a Web sphere/Web archive approach.

The September 11 Web Archive consists of Web sites related to the airliner attacks in the U.S. on September 11, 2001, archived between September 11, 2001 and December 1, 2001. During this period, Steve Schneider and I worked with the Pew Internet and American Life Project, the U.S. Library of Congress, the Internet Archive and volunteers from around the world to identify and archive URLs that were likely to be relevant to the question of how Web site producers were reacting to the events of September 11. Twelve basic categories of site producers were identified that were expected to be responding to the attacks on the Web. The findings for the studies on personal expression summarized below were based on an examination of Web sites produced by nine of these: news organizations such as CNN, the New York Times and Salon.com; federal, state and local government entities; corporations and other commercial organizations; advocacy groups; religious groups, including denominations and congregations; individuals acting on their own behalf; educational institutions; portals, and charity and relief organizations.

To build the archive, systematic searches were conducted for URLs produced by these sets of actors, and links to other URLs were followed to find more sites with relevant content. In most cases, the salient feature of these sites was content referring to the attacks and/or their aftermath. In some cases, the absence or removal of such content was salient. These collection efforts identified nearly 29,000 distinct URLs. Each site was archived on a daily basis from initial identification to the end of the collection period. The objective of the archiving activity was to

preserve not only the bits and the content, but also the experiential dimensions of this rapidly emerging Web sphere. By capturing pages and sites in their hyperlinked context, the archiving process preserved not just the collection of Web pages, but an interlinked Web sphere, characterized and bounded by a shared object orientation or reference point, in this case, the September 11 attacks.

The first study on expression as one form of sociopolitical action (Foot and Schneider, In press), included analysis of the types of site producers that enabled Web users to contribute personal expression or access expression posted by others on their sites, and whether mechanisms of expression were produced autonomously on a site (onsite), or jointly across sites (co-produced). For this analysis, a sample of 247 sites was generated from the September 11 Web archive. The sampling strategy, designed to include a broad representation of site producers, and to focus on those sites that were captured closest to September 11, yielded a sample of three "impressions" or site captures of the different Web sites. A preliminary analysis of the site pages had eliminated those without content relevant to the September 11 events, as well as those not captured in a readable format by the archiving tools. The refined sample of Web sites was then closely examined by trained observers for the range of social and political actions made possible by site producers, including personal expression.

Not surprisingly, we found that personal Web sites produced by individuals were most likely to both give Web users access to others' expression (typically the personal reactions of the site producers), and enable them to provide their own expression to the site (see Table 1). More interesting was that at least a quarter of all sites and over forty percent of sites in our sample

produced by news organizations, government entities, charities and educational institutions enabled the provision of personal expression by Web users. We interpreted these findings as suggesting an increased willingness in facilitating multi-voiced discourse if not dialogue on the part of site producers who might normally have vested interests in maintaining content control and a more singular voice on their sites.

Table 1: Percent of Sites, by Producer Type Enabling Expression

	Type of Site Producer									
Action enabled	**News**	**Government**	**Business**	**Charity**	**Advocacy**	**Religious**	**Personal**	**Educational**	**Portal**	**All Sites**
Get Expression	54%	16%	47%	50%	55%	53%	86%	76%	37%	55%
Provide Expression	50%	42%	23%	44%	32%	26%	69%	47%	26%	44%
Number of Sites	24	38	30	18	22	19	59	17	19	247

In this study we also took note of the mode of production employed in enabling expression. We defined an autonomous or on-site mode of production as one in which the site producer provides the content directly. In contrast, a joint or co-production mode is evidenced when a site producer links to another site to facilitate the user action, in this case accessing or providing personal expression. As Table 2 illustrates, many site producers combined these modes of production, providing some of the content themselves and linking to another site for additional content or functionality.

Table 2: Mode of Production in Enabling Expression		
	Action Enabled	
Mode of Production	Get Expression	Provide Expression
On-site	80%	75%
On-site & Co-produced	13%	15%
Co-produced only	8%	11%
Based on analysis of 247 sites		

Of those site producers whose sites enabled access to or provision of personal expression by Web users, a strong majority did so autonomously. Most site producers who engaged in co-production in enabling expression did so in addition to providing on-site access to expression and/or mechanisms for users to express themselves.

In the second study of online expression my co-author and I employed textual analysis to explore the particular forms of expression manifested on the Web (Siegl and Foot, 2004). This study shed light on the types of public expression evoked by personal or mediated exposure to a crisis, and posted on the Internet, and served as a case study in collective mourning on the Internet. The research questions guiding this analysis included: 1.) what

kinds of expression were posted on the Web after 9/11? and 2.) how do these forms of online expression compare with public mourning and bereavement? Using the same sample of daily impressions of 247 sites from the September 11 Web Archive described above, we identified 84 sites that enabled site visitors to post their own textual expression and/or access the textual expression of others. As in the previous study, these Web sites represented a broad cross-section of Web site producers including personal or individual sites, charity or civic organizations, businesses, and governments, as well as Web sites constructed for the sole purpose of memorializing the attacks. Due to the large variance in the nature of the Web sites, it was necessary to standardize the portion of expression observed on each site. This was accomplished by analyzing the first five discrete units of textual expression; a discrete unit was defined as a temporally bounded entry posted to the Web by an author.

Through close readings of the selected units of textual expression from archived impressions of each site, we identified nine forms of expression manifested on the Web in the three weeks following September 11 (9/11/01-10/2/01), and noted changes in dominant forms of expression during that period that deserved further study. We then compared the post-9/11 Web expression with emotional phases identified in the literature on public mourning and bereavement, such as shock, anger, and grief. We demonstrated that post-September 11 Web expression included more than these emotions, suggesting that the functions of the Web-based post-9/11 expression went beyond public mourning and bereavement and included attempts at analysis, sense-making, and advocacy. We concluded by arguing that the broader range of expression on the Web after September 11

(in contrast with expression documented from offline/non-Web contexts in the public mourning and the bereavement literature), is at least partially due to characteristics of the Web and processes/practices of Web production that distinguish it from traditional broadcast and print media.

The September 11 Web Archive and the two studies on post-9/11 online expression described above illustrate the usefulness of thematic Web archives and Web sphere analysis in facilitating investigations of some kinds of cybercultural phenomena. The demarcation of a Web sphere requires systematic identification of Web site producers as well as particular sites, which in turn creates a strong base for analyzing patterns and modes of Web (co)production as demonstrated in the first study. An archive of the Web sphere, collected at regular intervals during a specific period, enables retrospective and developmental analyses of many aspects of online relations, as well as the means and artifacts of cybercultural production and exchange. Web sphere analysis can function as a framework for research on sociocultural phenomena manifested in Web texts, features, or links, at a micro or macro level, and employing a diverse range of methods. As scholars of cyberculture undertake broader and deeper studies of Web form and content, as well as processes and patterns of production, distribution, usage and interpretation of Web-based phenomena, archive-based Web sphere analysis may provide a helpful foundation.

References

Arms, W., Adkins, R., Ammen, C. and Hayes, A. (2001) 'Collecting and Preserving the Web: The MINERVA Prototype', in *RLG DigiNews,* 5 (2): http://www.rlg.org/preserv/diginews/diginews5-2.html

Foot, K. A. (2002) 'Pursuing an evolving object: Object formation and identification in a conflict monitoring network', *Mind, Culture and Activity,* 9(2)**:** 132-149.

Foot, K. A. and Schneider, S. M. (2002) 'Online action in campaign 2000: An exploratory analysis of the U.S. political Web sphere', *Journal of Broadcasting & Electronic Media,* 46(2)**:** 222-244.

Foot, K. A. and Schneider, S. M. (In press) 'Online structure for civic engagement in the post-9/11 Web sphere', *Electronic Journal of Communication*.

Foot, K. A., Schneider, S. M., Dougherty, M., Xenos, M. and Larsen, E. (2003) 'Analyzing linking practices: Candidate sites in the 2002 U.S. electoral Web sphere', *Journal of Computer-Mediated Communication,* 8(4). http://www.ascusc.org/jcmc/vol8/issue4/foot.html

Hecht, M. L., Corman, S. R. and Miller-Rassulo, M. (1993) 'An evaluation of the drug resistance project: A comparison of film versus live performance media', *Health Communication,* 5(2)**:** 75-88.

Hine, C. (2000) *Virtual ethnography,* Thousand Oaks, CA: Sage.

Howard, P. (2002) 'Network ethnography and hypermedia organization: New organizations, new media, new myths', *New Media & Society,* 4(4)**:** 550-574.

Kahle, B. (1997) 'Preserving the Internet', *Scientific American,* 276(3)**:** 82-83.

Leont'ev, A. N. (1978) *Activity, consciousness, and personality,* Englewood Cliffs: Prentice-Hall.

Lindlof, T. R. and Shatzer, M. J. (1998) 'Media ethnography in virtual space: Strategies, limits, and possibilities', *Journal of Broadcasting and Electronic Media,* 42(2)**:** 170-189.

Lyman, P. (2002) 'Archiving the World Wide Web', in *Building a National Strategy for Digital Preservation,* Council on Library and Information Resources.

Lyman, P. and Kahle, B. (1998) 'Archiving Digital Cultural Artifacts: Organizing an Agenda for Action', *D-Lib Magazine*. http://www.dlib.org/dlib/july98/07lyman.html

Schneider, S. M. and Foot, K. A. (2004) 'The Web as an object of study', *New Media & Society,* 6(1)**:** 114-122.

Schneider, S. M. and Foot, K. A. (2005) 'Web sphere analysis: An approach to studying online action', In *Virtual Methods: Issues in Social Research on the Internet,* Hine, C. (ed.), Oxford: Berg.

Siegl, E. and Foot, K. A. (2004) 'Expression in the post-September 11th Web sphere', *Electronic Journal of Communication,* 14(1-2). http://www.cios.org/www/ejc/v141toc.htm

Stowkowski, P. A. (2002) 'Languages of place and discourses of power: Constructing new senses of place', *Journal of Leisure Research,* 34(4)**:** 368-382.

Taylor, J. R. and Every, E. J. v. (2000) *The emergent organization: Communication as its site and surface,* Mahwah, New Jersey: Lawrence Erlbaum Associates.