

Julien Masanès

Web Archiving

With 28 Figures and 6 Tables

Author

Julien Masanès
European Web Archive
25 Rue des Envierges
75020 Paris
France
julien@europarchive.org

ACM Computing Classification (1998): H.3, H.4, I.7, K.4
Library of Congress Control Number: 2006930407

ISBN-10 3-540-23338-5 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-23338-1 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable for prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media
springer.com
© Springer-Verlag Berlin Heidelberg 2006

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting by the author and SPi

Cover design: KünkelLopka, Heidelberg

Printed on acid-free paper SPIN: 11307549/45/2145/SPi 5 4 3 2 1 0

2 Web Use and Web Studies

Steve Jones and Camille Johnson

University of Illinois at Chicago
sjones@uic.edu

2.1 Summary

In 2002, the Online Computer Library Center (OCLC) estimated that over three million publicly accessible websites existed on the World Wide Web (O'Neill et al. 2003, para. 9). In the material world of information, this number would approximate from 14 to 28 million books, matching or exceeding the number of volumes held by most of the world's largest libraries. This overwhelming collection of information represents an endless, accessible supply of both verbal and visual data for scholars interested in studying online activity. And although the sheer volume of available resources on the Web has presented the challenge of choosing *what* to study and *how* to study, the Web has proven even more daunting. In this chapter, we will provide an overview of methodological approaches researchers have used to study the Web.

The goal in doing so is not to provide an exhaustive categorization of methodologies. Instead, it is our hope that by understanding the methods used for studying the Web and studying Web use, those who seek to archive and preserve the Web can better understand the needs of the research community.

The Web consists of an immense variety of types of materials whose variety is best understood along two dimensions. First, the Web itself is a medium, and not simply content. More to the point, it is both a medium that conveys content via numerous protocols (such as HTTP), and it is also a "container" for content, one which further "shapes" content and also presents it to viewers. However, the presentation of content is at the viewer's discretion and is further shaped by the viewing tools used (browsers and other applications). In other words, though the content may be the same Web page, two viewers using different browsers or using different browser settings may ultimately see different pages. Second, and further unlike the

30 Steve Jones and Gummie Pemberton

world of analog materials, the very definition of media for storing the Web is open to question. Websites can be locally stored or cached, may be mirrored, may be dynamic, or may be virtually ephemeral, as is the case with Web cams.

Internet research, in general, and Web research, in particular, is populated by a diverse group of scholars. Their various disciplines include linguistics, journalism, political science, business administration, geography, advertising, communication, and the arts, to name only a few. Because we are dealing with such a wide array of academic traditions, the types of materials with which they work are very diverse, and the working definitions of research methods vary as well. What counts as "ethnographic" analysis of Web texts for a marketing scholar may be interpreted as a qualitative content analysis by a communication scholar. Therefore, another goal of this chapter is not to provide rigid Web content categories or blueprints for each methodological approach as they apply to Web studies, but rather to present a range of interpretations and applications of these methods as they have proven most useful to researchers of the Web.

2.2 Content Analysis

Content analysis is one of the more common methods for studying the Web. As with other media, a researcher who uses content analysis codes Web content, either written text or images, based on particular criteria and places them within relevant categories or themes; in other words, it is a survey of Web content, rather than Web users. Within Web studies, content analysis has primarily been used as a comparative tool, allowing the researcher to make meaningful comparisons of content between similar Web texts. A study of antiglobalization organizations' websites used content analysis to determine whether cohesion in message and purpose existed among the sites (Van Aelst and Walgrave 2002). Content was coded and sorted based on what appeared to be its primary function within four areas: to provide information about the organization, to provide information about antiglobalization issues, to foster interactivity with the group organizers and other members, and to foster mobilization on behalf of the cause, such as donating money signing an online petition. Through their analysis, Van Aelst and Walgrave were able to conclude that indeed, antiglobalizations organizations online were generally using their websites in similar ways to inform and involve their memberships.

Another comparative study analyzed the content of radio station websites to determine how the commercial radio industry was responding to

the availability of the World Wide Web for promoting their stations. This assessment was made by analyzing the types of user-based information being provided on radio stations' websites (e.g., traffic maps and program logs), their use of websites as tools for promoting their station (e.g., contest information and DJ bios), and their incorporation of interactive features (e.g., e-mail addresses for the stations' staff members and listener surveys) (Pitts and Harms 2003). For radio stations, this Web study provided a valuable resource for evaluating their current and potential uses of the Web.

Content analysis has also been used to study the impact of institutional policy changes on related websites. Educational policies, for example, have placed increasing importance on the incorporation of information and communications technology (ICT) into secondary schools' institutional missions. One study chose to evaluate the progress being made toward achieving this goal as outlined by the National Grid for Learning in the UK through a content analysis of 150 secondary school websites (Hesketh and Selwyn 1999). Images and written texts were coded, identifying the websites within one of the following attitude categories: Proactive, pupil-centered, or reactive. Through these categories and an analysis of the schools' profiles, the researchers recognized a correlation between a commitment to ICT integration and institutional capital; the more economic and social capital a school had, the more likely that their website's content would reflect a positive, proactive attitude toward the incorporation of ICT in their mission.

The US Federal Trade Commission has employed content analysis to evaluate the implementation of online privacy and information security notices on commercial websites (Milne and Culnan 2002). Their study comprised a longitudinal analysis of four Web "surveys" spanning from 1998 to 2001. Criteria such as the inclusion or exclusion of notices regarding the disclosure of visitor information with third parties and use of "cookies," collection of user information beyond e-mail addresses, and the use of information security on the sites were used to determine whether the websites met the requirements of fair information practices. The results from the 2001 analysis were also compared with previous years and indicated that among commercial websites that collect personal information from their users there has been an increase in the implementation of fair information practices (p. 355).

2.3 Surveys

Web surveys can be defined in two distinct ways: those which are surveys *on* the Web, and those which are surveys *of* the Web. The former type appears to be the most common, with researchers posting surveys on websites in order to access their population of interest. In most cases, their purpose is to collect information about people's uses of the Internet,¹ such as the Graphic, Visualization, and Usability Center's (GVU) World Wide Web User Survey (1998). The GVV project is an ongoing study of Internet use that began in 1994. The project recruits participants through Internet-related newsgroup postings, banner advertisements on news media and search engine websites, and announcement in offline media such as magazines and newspapers. Several focused surveys have been generated by the GVV including a general demographics survey, a technology demographics survey which collected info such as connection speed of users and Internet browser used, a survey of users' attitudes toward online privacy and security, a general survey of computer, Web and Internet use, as well as a survey of online product searching and purchasing activities. Sample sizes typically topped 1,000, with the general demographics survey surpassing 5,000 participants.

Organizations have also found Web surveys a useful tool for assessing the quality of user experiences using related websites or Web functions. One such study conducted a survey of approximately 450 museum website visitors, collecting both demographic information and responses to the quality and usability of the sites (Sarraf 1999). Many museums have begun to provide websites as a means for supplementing information about their collections as well as creating an interactive, accessible relationship with potential and past visitors (p. 1). Qualitative and quantitative responses were solicited, providing rich, detailed feedback for the museums involved. Carswell and Venkatesh (2002) were able to use a Web-based survey to solicit evaluative responses from over 500 graduate students who

¹ With Internet access reaching 59% among US adults in 2002 (Spooner, 2003), researchers in the US are also beginning to consider Web-based surveys a practical substitute for traditional types of survey administration, such as paper and telephone, on topics other than Internet uses and attitudes. Researchers interested in collecting timely responses to current events seem particularly interested in Web-based surveys, exemplified by studies the impact of the terrorist attacks in New York City on September 11, 2001 on participants (e.g., Lee et al. 2003). Just nine days after the attacks, one team of researchers was able to distribute a survey via the Web analyzing the psychological responses of Americans to this traumatic event (Silver et al. 2002).

had participated in asynchronous online courses. They found partial support for their hypotheses that acceptance of and future intent to use technology in an asynchronous online course would be positively influenced by the student's attitudes and perception of the technology. In all cases, including the GVV project, the benefit of using a Web-based survey was clear: the populations of interest were Web-users, and the data collected related directly to their Web activities.

2.4 Rhetorical Analysis

Another subset of Web research looks critically at Web texts to identify persuasive strategies through rhetorical analysis. As defined by Warnick (1998)

...a rhetorical critical method considers how the text gives presence to some elements as opposed to others, how narrative constructions configure browser experience in certain ways, and how the discourse plays to the predispositions and habits of mind of its audience. (p. 309)

In her study of political websites during the 1996 US presidential campaign, Warnick specifically analyzed the rhetorical strategies of parodic websites – sites which mimicked the design and content of legitimate campaign websites, possibly to erode the credibility of the sites mimicked and certainly to provide comic relief (p. 308). She concludes that the parodic sites relied on popular narratives of distrust in government, the simulation of political participation through interactive features such as petitions which were never delivered to their intended party, and innuendo regarding candidates' political and criminal pasts to persuade their readers to adopt their rhetorical vision. Ultimately, Warnick finds the strategies hypocritical, as they include many of the same unethical behaviors purportedly used by the candidates in question.

Rhetorical analysis of websites has been used to identify rhetorical communities, those groups which use the Web to create their own worldview which may or may not agree with popular narratives. Kroeber (2001) discusses how a community of websites created by feminist mothers presented a platform for participants to actively resist negative, feminist conceptualizations of motherhood. The women were able to renegotiate feminism through the written text and images found on their websites, arguing for an understanding of motherhood as an empowering, even integral, part of being a feminist.

Another study looked at the use of websites for persuasive purposes by hate groups such as the Knights of the Ku Klux Klan, Nation of Islam, and the Neo-Nazi organization National Alliance (Duffy 2003). In this case, Duffy used a rhetorical critical approach called Fantasy Theme analysis, which relies on Symbolic Convergence Theory. As defined by Duffy, Symbolic Convergence Theory is “a general theory of rhetoric in which groups create and share fantasies about the group and outside groups and thereby build a shared identity” (p. 293).² Through an analysis of selected written texts on each group’s website, the author was able to determine several rhetorical narratives, including “the plea for fairness and justice,” “the natural order and the resurrection of the people,” and “the original people of the Earth are called to give new meaning to race” (pp. 295–305).

2.5 Discourse Analysis

McQuail (2000) describes discourse analysis as applying “to all forms of language use and textual forms,” relying on the idea that “communication occurs by way of forms of ‘text and talk’, adapted to particular social locations, topics and kinds of participants” (p. 494). Discourse analysis of webtexts considers the socio-cultural positioning of websites; their construction of localized meaning through verbal and visual elements. Again, hate group Web texts surface as Billig’s (2001) study explores the language conventions of humor as it is subverted by Ku Klux Klan “joke” websites. He found that disclaimers were used to warn against the interpretation of racist jokes on the sites as dictating violent action that the content was simply meant as a joke for those with “a sense of humor” (p. 274). However, the author concluded that much of the content was not presented as a joke, but as fact — a meta-discourse which he states “denies that the joke is a joke” (p. 278). Racist texts and images, both addressed to the fellow racist readers of the “jokes” and the ethnic targets of the jokes, provide evidence of a more diabolical discourse, one which Billig’s likens to the social interactions of a lynch mob (p. 287).

Macrolevel discourse has also been analyzed on the Web, concentrating on websites which represent the voice of nation-states. In Purcell and Kodras’s (2001) study of the Slovenian government’s website, they approach

² Foss (1996) describes Symbolic Convergence Theory as being based on two primary claims: that “communication creates reality,” and that “symbols not only create reality for individuals but that individuals’ meanings for symbols can converge to create a shared reality for participants” (p. 122).

its texts as a response to perceived inaccuracies in the representation of Slovenia as a Balkan state; a region plagued by civil war and therefore framed as undesirable by Western countries for tourism and other commercial investments. Purcell and Kodras observe that Slovenia attempts to reinvent its identity through the texts and images on its government website, a rhetorical strategy. But they conclude that the site is ultimately part of an “an ongoing effort to negotiate [Slovenia’s] position” within the “global discourse” (p. 364).

Chandler (1998) has demonstrated that, in addition to institutional websites, personal homepages may serve as a form of discourse. In his analysis of Web texts created by teenagers in Wales, and through interviews with the site creators, Chandler explores the many ways the teens conceptualize their audience and negotiate the boundary between the public and private spaces of their lives. Many of the young Web authors stated that their sites represented an attempt to communicate with visitors who might share the same interests, while others described it as something they had created just for themselves, while still others simply felt motivated to be a part of the “vast” Internet and to share their thoughts about life with others (p. 12).

2.6 Visual Analysis

Due to the multimedia capabilities of the World Wide Web, some researchers have chosen to supplement or circumvent more traditional analyses of verbal Web texts with analyses of visual texts. Website creators have employed many creative uses of graphics and images on websites, with few garnering as much browser and media attention as JenniCam.org. This personal website features a 24 h, camera’s eye view of Web author Jenni Ringley’s bedroom. Jimroglou (1999) uses the JenniCam site as a case study for exploring Haraway’s (1991) concept of the cyborg, “a hybrid of machine and organism” (p. 149). As with the cyborg, Jimroglou observes that the boundaries between Jenni’s body and technology become blurred through her camera presence online, using as an example the image of Jenni sitting in front of her computer monitor, the site of her Web cam: “The image of Jenni at her computer becomes an icon for that fusion, with her flesh melting into her keyboard” (p. 441). Feminist film theory is used to explore the meaning of JenniCam as a gendered, visual subject. Ringley’s dedication to providing a glimpse into her real life includes leaving her Web cam on even when nude or engaging in sexual activity. Jimroglou recognizes the unique problem presented by these images: Jenni succeeds in breaking traditional boundaries of public and private spaces assigned to

the presentation of a woman's body while also evoking criticism from some feminists who feel she's fallen into the trap of female body objectification rampant on the Web. In either case, "Jenni's body functions as the locus of meaning, as the site of plentitude, as the root of unified meaning of JenniCam" (p. 449).

Other researchers have employed visual analysis to evaluate news media's use of Internet for presenting and distributing images. Fair and Parks (2001) present a critical analysis of images of the 1994 Rwandan genocide crisis made available by US news media via TV and the Web. Ground coverage of the crisis, in which cameras primarily captured images of frightened refugees being assisted by white, western relief workers, as well as aerial satellite images of people fleeing Rwanda en masse, are argued to have disempowered the victims of the crisis, oversimplifying the political context of the event, and distancing the viewer from the people and issues involved. Fair and Parks add that use of websites to present and distribute the Rwandan images is part of a larger trend in the domination of ownership of visual technologies by western news media, the content being produced chiefly for the consumption of western audiences, the primary users of the Internet. Fair and Parks conclude that image choices made by news media in their coverage of the Rwandan crisis, in conjunction with the mass distribution of the images through western media technologies, worked together to reinforce existing American "dislocation" from the culture and politics of Africa (p. 42).

The two studies reviewed above pay exclusive attention to images used on websites in order to render their interpretations. But rather than being mutually exclusive, the analyses of visual and verbal texts are typically integrated by researchers, as with Hesketh and Selwyn's (1999) study of the construction of school identities through images and texts on websites or Warnick's (1998) study of parodic campaign websites and their use of both verbal and nonverbal texts to advance particular political agendas. The visual design of websites, including the use and placements of images and texts on the page, provides another example of this integration of methods (Rivett 2000). A case study of the Volkswagen "Newbeetle" website discussed layout choices, such as the predominant splash image featured on the site's homepage, which the author compares to the conventions of off-line magazine design. The site was also said to communicate the company's online "visual identity" through the use of a white background (p. 50). The visual analysis of the "Newbeetle" website complements a textual analysis, revealing a narrative of "alien invasion" woven throughout its pages. The VW site was then compared to the visual layout of Daniel Chandler's academic site. The latter was described as being dominated by text, emphasizing its purpose as a source of information; a ring-binder

graphic was used as the background image for the site, reinforcing its academic nature.

2.7 Ethnography

Ethnography is most easily understood as “a description and interpretation of a cultural or social group or system” (Creswell 1998; p. 58). Defining the case in ethnographic research has become a more complex process as it is performed within the study of Internet and Web cultures. The praxis of pre-Internet ethnography has been the study of single-sited, geographically centralized cultures. However, online unity and offline physical dispersion of participants in Internet cultures has required a more loosely bounded, multisited approach to ethnography (Howard 2002). The notion of the multisited community may also need to be applied to the “locations” of a community’s participants online as well; providing a “thick description” (Geertz 1973) of an online community requires a more varied approach than simply observing group discourse within a chatroom or analyzing the contents of community websites. Instead, some Web ethnographers have argued that a more holistic approach is required, one in which all sites of participation and experience of community members are explored by the researcher, both on and offline (Howard 2002; Miller and Slater 2000).

Miller and Slater (2000) provide a useful example for performing a multisited model of Internet ethnography in their study of Trinidadians and Internet. The researchers visited chat rooms and websites, as well as people’s homes and cybercafes, in order to capture the way Internet has intersected with the social, political, economic, and religious lives of “Trinis.” Each site is described as providing a unique perspective on the culture; where chatrooms were thought to give insight into *being* “Trini” online, websites were conceived as *representations* of Trinidad online (p. 103). Miller and Slater use a qualitative content analysis to discuss the websites, pointing out the weaving together of national symbols, such as flag and map images, with the personal identities of the Web authors throughout the sites analyzed. They also discuss the reflection of political agendas in the websites, with Trinidad’s most profitable tourist attraction, the Carnival festival, being a frequent focus of pride and discussion on the sites.

Researchers interested in online fan communities have found ethnography a useful approach to their studies. Bloustien (2002) included analyses of fan websites and related discussion forums in her depiction of fandom and the television show *Buffy the Vampire Slayer*. From her analyses of Web texts, as well as data collected from in-person interviews and tours of

the fans' rooms, she was able to draw conclusions about how *Buffy* fans, mostly teenagers, draw on the fantasy and magic themes of the show to maintain the imagination and "play" of youth, while entering into the more serious and "terrifying possibility of moral ambiguity" found in adulthood (p. 440).

2.8 Network Analysis

According to Garton et al. (1997), "a social network is a set of people (or organizations or other social entities) connected by a set of social relationships, such as friendship, co-working or information exchange" (para. 2). When these social relationships are established, maintained, and/or constructed through the use or creation of websites, hyperlinks provide a means for studying the patterns of association between network members. Park (2003) identifies this type of analysis in Web research as hyperlink analysis, in which websites are conceptualized as representing individuals, groups, organizations, and nation-states, and hyperlinks between the sites as representing "relational connection[s]" (pp. 50–51). For instance, Halavais (2000) studied the hyperlinks of websites in order to better understand whether the Web is truly fostering a "World Wide" network of sites, or whether national borders are being adhered to online. He looked at 4,000 websites, examining their links to pages external to themselves, followed by an identification of the national host of the websites linked to. Through this process, Halavais was able to conclude that in fact most websites in his sample linked primarily to sites within their respective national cultures. Yet in comparison with other types of information technologies, such as ground mail and television, Halavais believed the Internet to be the most "internationalized," with more references to and from the US over international borders occurring than with other media (p. 23).

Organization-level network analysis was used in tandem with content analysis in the study of antiglobalization websites, mentioned previously, in order to understand to what degree these sites were integrated (Van Aelst and Walgrave 2002). Links between the 17 sites were analyzed using network visualization software, Pajek (<http://vlado.fmf.uni-lj.si/-pub/networks/pajek/>), which created a graphic map of the links between the websites. Highly connected and referenced sites were easily identified by the clustering of links to and from their location, while less integrated sites appeared as isolated outliers with only one or two links connecting them to the rest of the network. Van Aelst and Walgrave found that the antiglobalization sites were fairly integrated, but warned that it is difficult to

assess the nature of the relationships between organizations based purely on the existence of links between sites.

In other cases, hyperlinks have been analyzed in order to uncover the social and informational networks of individuals online. Most Web browsers, such as Netscape Navigator or Internet Explorer, include a "History" function which saves the URLs accessed by a person while visiting sites on the Web. Tauscher and Greenberg (1997) used this function to analyze the browsing behaviors of 23 individuals over a six week period, looking for patterns in webpage visitation. They found that nearly two-thirds of subjects' webpage visits were ones which had been previously viewed (p. 112). In addition, the number of pages revisited frequently by participants was relatively small, further representing rather limited personal site networks among participants. Tauscher and Greenberg speculate that the "Back" function on Web browsers, which constituted 30% of their subjects' navigation actions, may be a contributing factor to Internet users' penchant for page revisitation on the Web (p. 131).

In addition to social networks the study of document networks has grown, largely due to the sheer size of the Web, which has meant that various forms of social scientific network analysis are possible. The analysis of document networks can provide opportunities for meta-analysis of content. Henzinger and Lawrence (2004) discuss methods of sampling Web pages "to automatically analyze a large sample of interests and activity in the world...by analyzing the link structure of the Web and how links accumulate over time" (p. 5186). Aizen et al. (2004) posit that "usage data at a high-traffic website can expose information about external events and surges in popularity that may not be accessible solely from analyses of content and link structure" (p. 5254).

2.9 Ethical Considerations

The importance of the need for Internet scholars to pursue research following ethical practices is clear, but the policies and means of ethical research practice are less clear. For scholars studying the Web, and particularly for those who mine archival Web-materials, ethical considerations will inevitably arise during the course of research. Some material found on the Web and that is being archived is confidential, inadvertently made available and then stored by search engines like Google. The ethical positions concerning use of private material publicly archived are numerous and beyond the scope of this essay. However, it is important to note that the thoroughness with which search engines scour the Web can lead to the archiving of material that users

had neither intended to make public nor to have archived. One, perhaps overly optimistic, estimate of the speed with which search engines index the Web noted that indexing and archiving by the likes of Google outpace the creation of new Web pages (Whelan 2004). It is likely that issues that have been long discussed by Internet Studies scholars (Ess 2002) and those studying Computer-Mediated Communication concerning public versus private data will be soon at the forefront among Web studies scholars and those involved in archiving Web and other electronic materials.

2.10 Conclusion

Clearly, Web studies have come into their own. No longer in the shadow of textual analyses of text-based communities, Web research is proving that concepts such as community, culture, behavior, and meaning construction can also be effectively examined on the World Wide Web. While the wealth and variety of Web research being conducted is promising, there are still many avenues of inquiry yet to be explored.

The legacy of textual analysis of computer-mediated communication is apparent in the heavy bias toward language related analyses in Web studies. Missing are studies which explore the Web as a multimedia-landscape, describing how images and sound are being used to circumvent verbal communication, to overcome language barriers in a global medium. With the exception of the mention of sound use on Trinidadian websites in Miller and Slater's (2000) ethnographic study, aural Web studies were completely absent from the existing body of Web research. Future projects might examine rhetorical uses of sound on the Web via analyses of commercial or political websites. Other multimedia-oriented studies could look at the ways sound and images are being used to construct identities of persons online, from individuals to organizations to national governments.

The challenges for preservation of Web content are discussed elsewhere in this volume. For Web studies it is imperative the challenges are overcome. To serve the growing number of researchers it is necessary to have Web content of all kinds, from all periods, easily accessible. Furthermore, it is necessary to be able not only to access content, but also to recreate the links within content and between sites. And, it is necessary to have available the browsers and other applications with which Web content is viewed to best understand the experience of the user. The challenges thus go beyond preservation of content and include preservation of structure of, and encounter with, content.

References

- Aizen, J., Huttenlocher, D., Kleinberg, J., & Novak, A. (2004). Traffic-based feedback on the Web. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(1), 5254–5260
- Billig, M. (2001). Humour and hatred: The racist jokes of the Ku Klux Klan. *Discourse & Society*, *12*(3), 267–289
- Bloustein, G. (2002). Fans with a lot at stake: Serious play and mimetic excess in *Buffy the Vampire Slayer*. *European Journal of Cultural Studies*, *5*(4), 427–449
- Carswell, A. D. & Venkatesh, V. (2002). Learner outcomes in an asynchronous distance education environment. *International Journal of Human-Computer Studies*, *56*, 475–494
- Chandler, D. & Roberts-Young, D. (1998). The construction of Identity in the Personal Homepages of Adolescents. Retrieved April 15, 2002 from <http://www.aber.ac.uk/media/Documents/short/strasbourg.html>
- Creswell, J. W. (1998). *Qualitative inquiry and research design*. Thousand Oaks, CA: Sage
- Duffy, M. E. (2003, July). Web of hate: A fantasy theme analysis of the rhetorical vision of hate groups online. *Journal of Communication Inquiry*, *27*(3), 291–312
- Ess, C. (2002). Ethical decision making and Internet research: Recommendations from the AoIR working committee. Association of Internet Researchers. Retrieved July 12, 2004, from <http://www.aoir.org/reports/ethics.pdf>
- Fair, J. E. & Parks, L. (2001). Africa on Camera: Television news coverage and aerial imaging of Rwandan refugees. *Africa Today*, *48*(2), 34–57
- Foss, S. K. (1996). *Rhetorical criticism: Exploration and practice*. Prospect Heights, IL: Waveland
- Garton, L., Haythornthwaite, C., & Wellman, B. (1997). Studying online social networks. *Journal of Computer-Mediated Communication*, *3*(1). Retrieved April 4, 2004 from <http://www.ascusc.org/jcmc/vol3/issue1/garton.htm>
- Geertz, C. (1973). *The interpretation of cultures*. New York: Basic Books
- Graphic, Visualization and Usability (GVU) Center (1998). 10th WWW User Survey. Atlanta, GA: Georgia Institute of Technology. Retrieved March 25, 2004 from http://www.cc.gatech.edu/gvu/user_surveys/survey-1998-10/.
- Halavais, A. (2000). National Borders on the World Wide Web. *New Media & Society*, *1*(3), 7–28
- Haraway, D. (1991). *Simians, cyborgs, and women: The reinvention of nature*. New York: Routledge
- Henzinger, M. & Lawrence, S. (2004). Extracting knowledge from the World Wide Web. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(1), 5186–5191
- Hesketh, A. J. & Selwyn, N. (1999). Surfing to school: The electronic reconstruction of institutional identity. *Oxford Review of Education*, *25*(4), 501–520

- Howard, P. N. (2002). Network Ethnography and the Hypermedia Organization: New Media, New Organizations, New Methods. *New Media and Society*, 4(4), 550–574
- Jimroglou, K. M. (1999). A Camera with a view: JenniCam, visual representation, and cyborgsubjectivity. *Information, Communication and Society*, 2(4), 439–453
- Kroeber, A. (2001). Postmodernism, Resistance, and Cyberspace: Making Rhetorical Spaces for Feminist Mothers on the Web. *Women's Studies in Communication*, 24(2), 218–240
- Lee, W., Hong, J., & Lee, S. (2003). Communicating with American consumers in the post 9/11 climate: An empirical investigation of consumer ethnocentrism in the United States. *International Journal of Advertising*, 22, 487–510
- McQuail, D. (2000). *McQuail's mass communication theory* (4th ed.). London, UK: Sage
- Miller, D. & Slater, D. (2000). *The Internet: An ethnographic approach*. Oxford, UK: Berg
- Milne, G. R. & Culnan, M. J. (2002). Using the content of online privacy notices to inform public policy: A longitudinal analysis of 1998–2001 US Web surveys. *The Information Society*, 18, 345–359
- O'Neill, E. T., Lavoie, B. F., & Bennett, R. (2003). Trends in the evolution of the public Web, 1998–2002. *D-Lib Magazine*, 9(4). Retrieved March 29, 2004 from <http://wcp.oclc.org/>.
- Park, H. W. (2003). Hyperlink network analysis: A new method for the study of social structure on the Web. *Convergence*, 25(1), 49–61
- Pitts, M. J. & Harms, R. (2003). Radio websites as promotional tools. *Journal of Radio Studies*, 10(2), 270–282
- Purcell, D. & Kodras, J. E. (2001). Information technologies and representational spaces at the outposts of the global political economy. *Information, Communication and Society*, 4(3), 341–369
- Rivett, M. (2000). Approaches to analyzing the Web text: A consideration of the Web site as an emergent cultural form. *Convergence*, 6(3), 34–56
- Sarraf, S. (1999). A survey of museums on the Web: Who uses museum Web-sites? *Curator*, 42(3), 231–243
- Silver, R. C., Holman, E. A., McIntosh, D. N., Poulin, M., & Gil-Rivas, V. (2002). Nationwide longitudinal study of psychological responses to September 11. *Journal of the American Medical Association*, 288(10), 1235–1244
- Spooner, T. (2002). Internet use by region in the United States. Pew Internet & American Life Project. Washington, DC. Retrieved March 25, 2004 from http://www.pewinternet.org/reports/pdfs/PIP_Regional_Report_Aug_2003.pdf
- Tauscher, L. & Greenberg, S. (1997). How people revisit Web pages: Empirical findings and implications for the design of history systems. *International Journal of Human-Computer Studies*, 47, 97–137

- Van Aelst, P. & Walgrave, S. (2002, December). New media, New movements? The role of the Internet in shaping the 'anti-globalization' movement. *Information Communication and Society*, 5(4), 465–493
- Warnick, B. (1998). Appearance or reality? Political parody on the Web in Campaign '96. *Critical Studies in Mass Communication*, 15, 306–324
- Whelan, D. (2004, 16). Google Me Not. *Forbes*, 174(3), 102–104