

Online Citations: A Retrospective of Emerging Norms in Early Publications

Jennifer J. McGee^{※1}

When I was using material from Usenet in my PhD. Dissertation in the early 1990s, I struggled with a variety of questions about the ethics of citing these materials. Was it violating the authors' privacy to cite such material with their names attached when they clearly had not intended it to be read outside of the original newsgroup? If I were to cite the materials with pseudonyms, would it deny the authors the right to be associated with their words, which were after all written and posted in a relatively public forum, capable of being archived and accessed by others for the indefinite future? These questions made quoting and citing my materials an especially murky issue for me; in addition, at a certain level I began to feel the whole problem was--in every sense of the word--academic. After all, as Gurak points out in her appendix to Persuasion and Privacy in Cyberspace, in most cases more people read the original material online than will ever read it in the scholarly book, article, or convention paper (139). However, the questions remained for me, as they do for all researchers who work with online sources. This paper is a look at how a variety of researchers answered these questions when studying the Internet was still a relatively recent phenomenon in the 1990s, some patterns to be found in the answers, and areas of anomaly and flux.

This paper surveys 26 different qualitative studies of on-line communication with an eye to how the researchers dealt with the tension between private and public spaces. The books and articles studied are from a variety of sources: academic journals, collections of essays, and some books aimed at a more popular or interdisciplinary audience. All the materials studied either look at a specific Internet forum (i.e. one mailing list, one Usenet newsgroup, a specific set of Web pages), or use concrete data from different fora to make a more general point about the Internet (for example, Clerc looks at fandom on mailing lists and Usenet newsgroups). Either way, the researchers had to make decisions about how to present and cite the data used.

Each study examined was analyzed on the basis of five questions:

1. Which aspect of the Internet was being studied?
2. Did the author choose to use direct quotations from the data chosen?
3. Did the author use the real names (or at least the real on-line "handles") of the sources of the data?
4. Does the author discuss whether permission to use their words was sought from the sources?
5. If the author uses direct quotations, are they cited in a bibliography or works cited section?

All researchers have their own idiosyncratic reasons for presenting and citing the data the way they do, but

※1 言語コミュニケーション学科

the answers to questions 2, 3 and 5 reveal researcher biases toward a public or private orientation. Avoiding direct quotes, using pseudonyms, and not citing the material all indicate a concern for the privacy of the individuals cited, whereas a willingness to use the exact words, to reveal the identity of the speaker, and to cite the material as if it were a discrete source and not collected data point to an assumption of the publicness of the discourse. In addition, as Laura Gurak notes, the "privacy" bias is also based in part on an assumption of the orality of the data--"the notion that Internet texts are not texts at all but rather written versions of spoken conversations" (138). Researchers who use real names and cite their material are approaching it as if it were written, and at least semi-published, public information. As will be mentioned later, the private/oral assumption tends to correlate with specific types of on-line communication, and the public/written assumption with other types.

Methods of Presenting and Citing Data

Nearly all of the studies examined used one of four different strategies for presenting and citing on-line data: avoiding quotations entirely, using quotations but with no names attached, using quotations with real names, but no bibliographic citation, and using quotations with real names and full bibliographic information.

No Quotations or Oblique Reference. Some studies described the subject of their on-line ethnographies without ever actually using specific data from the group studied. Researchers using this technique referred to historical events or norms of the group without ever actually quoting members of the group. Baym and Anderson's descriptions of Usenet newsgroups, for example, discuss ways of interacting on their respective newsgroups without ever giving specific quotations from the groups. Similarly, Camp's essay on *Systers*, a mailing list for women in computer science, describes the culture of the list without ever quoting directly from it.

A similar tactic is to use material not from the group directly, but from indirect sources. For example, Ito's discussion of MUDs cites no material from the MUDs being studied; instead, she quotes statements taken from interviews done face to face, on the MUD, and via email. Kendall mainly uses the same strategy in "MUDder? I Hardly Know Her!," as does Turkle in her study of on-line personae, Life On the Screen, McRae in her discussion of MUDs, and David Shaw in his essay on gay men and IRC. A slightly different approach is Reid's analysis of MUDs, in which she quotes from neither the MUD nor personal interviews, but from Usenet newsgroups on MUDding, working under the assumption (a common one, as will be discussed later) that Usenet newsgroups are more public and published than a MUD.

In the cases other than Reid's, it is often unclear whether the material being used in these studies was collected from face to face interviews, on-line real time interviews, or personal email interviews; all three options are similar to face to face interviews in that the subject is (one assumes) fully informed of the purpose of the interview and aware that his/her words may be quoted by the researcher. Researchers using interviews in this way also face a decision on whether to use the actual real or screen names of the people interviewed. Kendall, McRae, and Ito never state explicitly whether the subjects' names were changed, but

Turkle changed all of her interviewee's names, and Shaw's interviewees were all given pseudonyms (pseudonicks?), a decision probably based in part on the especially personal nature of the forum.

Researchers using these indirect approaches to these sources face the difficulty of having to evoke and describe a group's norms without giving any clear, direct examples of the communication being studied. The researchers who choose to use no quotations at all have a particularly hard task. The three mentioned above succeed because it is apparent that each author is well-versed in the subject. Anderson, for example, does an overview of a variety of newsgroups that reveals her competence on the topic, and Baym and Camp have both a level of detail and a sense of history of their respective groups that reveals intensive research. This approach would not work as well for a researcher who had watched a group for shorter periods of time.

Direct Quotations and Pseudonyms

A second approach is to take direct quotations from the group being studied, but to delete the names or nicknames of the subjects. A few of the many studies choosing the approach include Carstarphen and Lambiase's study of a student mailing list, McLaughlin, Osborne and Ellison's ethnographic look at a Web-based message board, Aycock and Buchignani's look at a specific controversy on Usenet, Odzer's popular press book about sex and anthropology on MUDs, and McLaughlin, Osborne, and Smith's examination of standards of conduct on Usenet newsgroups. Kendall occasionally uses logs from a MUD in her discussion of MUDDing, and uses pseudonyms when she does. Clerc's look at Internet-based fandom examines Usenet newsgroups and mailing lists, supplemented with email interviews, and quotes from all three with no attributions. By stripping the names or user names from the quotations, the authors cited above have access to concrete data that supports their assertions about on-line groups, but still leave the statements essentially private and untraceable.

Quotes and Real Names, Uncited

A relatively uncommon method of quotation and citation is to use direct quotations without removing the subject's name or user name, but also without actually citing the data as a source in the bibliography. Stacy Horn, the founder of the New York virtual salon Echo, uses this technique in her book Cyberville. Schmitz's examination of PEN, Santa Monica's public bulletin board, also uses this method, as do two studies of Usenet, Benson's "Rhetoric, Civility, and Community," and Mitra's "Virtual Commonality." This style seems to be used fairly rarely, however, because researchers who are willing to print both direct quotations and use actual names are generally willing to go the extra step and actually cite the source in a bibliography.

Direct Quotation and Full Citation

The final possibility is for the researcher to treat the material from the group being studied as fully public and published and to actually cite it in a bibliography. Most authors willing to use real names give full citations for material, as in Watson's examination of posts to the phish.net web discussion group, Knapp and Tepper's essays on Usenet, or Zickmund's tour of racist and homophobic Web sites. Mitra's analysis of Indian Web pages goes even further: in using direct quotations from the Web pages under examination, Mitra quotes one poster's email address and cites that email address in the bibliography, a practice that seems to me

rather like putting an author's phone number next to his or her name in a bibliography.

All of the authors sampled use one of the four styles listed above, with one exception: Laura Gurak's Persuasion and Privacy in Cyberspace. Gurak uses pseudonyms but still cites the material in her endnotes, thus falling outside all the four categories. Gurak's own reason given for this is her attempt to balance privacy with author copyright in an area still under some contention; this level of balance means her work defies categorization.

Question #4, on whether permission was obtained from the subjects of the study, proved frustrating. Only Turkle explicitly discusses whether permission was sought, although many authors imply that it was or was not (for example, Gurak's discussion of method implies she did not, and it seems unlikely that Zickmund asked permission of the racist Web site developers). The authors who chose to use no direct quotations avoid the question entirely; I infer that no permission was sought. Those who use email interviews imply that permission was gained to use this information, but no one explicitly addresses the issue.

Conclusions

My overview of on-line studies revealed both some general patterns in citation and some exceptions to these patterns, which is not surprising considering both the youth of the area of study at the time, and the breadth and interdisciplinary nature of the articles. Online ethnographies and rhetorical criticisms had had little time to develop norms of citation and come from a variety of different academic backgrounds, making consensus potentially slow. Even now, consensus is not clear about how to properly use online sources, although more and more they are being considered public works. However, even back in the early 1990s there were clear emergent norms about quoting and citing material from online sources.

As has been pointed out, different online sources are generally seen as being more or less oral versus written. IRCs and MUDs are closer to spoken conversation in that they are in real time and are generally more informal. IRC and MUD communication is more ephemeral than Usenet posts, email, or Web pages--if no one logs the interaction, it is irretrievable, like most informal oral conversation. And IRC and MUD conversations both feel more private than other fora as well, because you can see who exactly is receiving your messages. Usenet posts, mailing lists, and web pages feel more like a broadcast--you send out your messages for all who happen upon them--and thus feel more public.

Reflecting this view, researchers who study real-time online conversations often take some steps to protect the privacy of their subjects and hide their true identity. Kendall, Shaw, and Turkle all explicitly state that the names of their subjects have been altered in their studies of IRC and MUDs. The other researchers, like Ito, McRae, Odzer, and Reid, do not clarify whether the names used are the subjects' real online names or not. Without knowing if the names are pseudonyms or not, it is hard for me to reach a definite conclusion about how carefully researchers guard the privacy of real-time online subjects, but roughly half of them are concerned enough to explicitly address the issue. Researchers studying IRC and MUDs are also much more likely to use indirect sources like email interviews or Usenet newsgroups--in part probably because then the

subjects can give informed consent, in part because the texts are less ephemeral than real-time chat. The only researcher I studied who actually cites real-time chat sources is Reid in her study of IRC, who footnotes the dates and times of her IRC logs being excerpted.

At the opposite extreme are Web pages, which are much more fixed, permanent, and public than IRC and MUDs. Because there is so little interaction between the creator of a Web page and the readers, it feels much more like a broadcast medium and thus published, public property. As a result, researchers who focus on Web pages usually show no qualms about quoting, naming, and citing material from Web pages. Zickmund, Mitra, and Watson all clearly cite their sources in their bibliography--Mitra, as mentioned before, even goes so far as to mention email addresses in a Web site's guest book. Grossman's book-length overview of the Internet relies mainly on essays and other material found on Web pages, all of which are clearly cited. The only work to study Web pages and not cite the material in their bibliography is the essay by McLaughlin, Osborne, and Ellison, and their work is focused on a Web-based message board rather than a standard Web page.

In between the relative privacy of real-time chat and the relative publicness of a Web page fall Usenet newsgroups and mailing lists. The four studies that examine material from mailing lists are all relatively privacy-oriented--Clerc and Carstarphen and Lambiase use direct quotes, but remove real names, while Camp uses no direct quotations and Gurak uses pseudonyms. Mailing lists are especially problematic because email is generally considered private, but mailing lists are semi-public: material on mailing lists is sent to a larger and relatively unknown audience. This ethical difficulty may explain why there were so few studies available about mailing lists in the 1990s.

The largest number of studies were of Usenet newsgroups, and researchers seem to be somewhat divided on the issue of the publicness of Usenet and how that plays out in their research. Of the twelve works that looked at Usenet, half of them use the more public forms of quotation and citation, and two thirds of those use full citations. The fully public and written assumptions were the ones I worked with in my dissertation as well. Half of the researchers, on the other hand, took steps to ensure the privacy of their subjects, either by using no quotations at all or by hiding the posters' real names. The publicness of Usenet is clearly still open to question; for now, each researcher chooses to balance between privacy and copyright as best they can to fulfill what is seen as the researcher's ethical obligation.

Although there are some emerging norms as to what is "oral" and private and what is "written" and public on the Internet, these norms have continued to fluctuate. This fluctuation is due to the quickly changing nature of the Internet, as well as the tendency of the different Internet fora to blur the dividing lines between them. For example, if you go to groups.google.com, you will find a searchable archive of a host of Usenet newsgroups. If you search this Web site, is the material you are using from Usenet newsgroups or from a Web site? Many mailing lists have "mirrored" Usenet newsgroups, which makes the line between private email and public communication harder to distinguish in those cases as well.

Finally, as the WWW becomes more and more sophisticated, the very ways we have of communicating on

it may evolve and continue to defy our ability to use them easily as material. The Internet changes so quickly that entire fora may fall out of use—for example, in the early 1990s, Usenet was a popular source of ethnographic material, but now Usenet itself has fallen out of favor among Internet users, most of whom now prefer to use message boards based on the World Wide Web. Still, just like Usenet, these message boards are an odd mix of private and public discourse. Javascripted chat is a mix of the private medium of chat and the public medium of a web page. The technology itself may keep the norms on citation from becoming overly stable for the indefinite future.

The conventions for quoting and citing Internet-based material are still not entirely settled, but there are some general patterns that are beginning to fall into shape, largely around the issue of the relative orality of the texts being studied. Real-time conversations, be they face to face or on-line, are generally considered private, asynchronous communications public. Deviations from these patterns reveal something about the assumptions the author makes about the nature of the specific texts being considered, and of the Internet in general. They also reveal a great deal about the complexity of ethical issues of publicness and privacy in the electronic world. However, as long as the Internet remains a source of such rich, easily collected (if not so easily sorted) ethnographic data, there will be people like myself and the other researchers cited in this study who are more than willing to grapple with the difficult issues of citation of electronic sources, no matter how they might evolve in the future.

WORKS CITED

- Anderson, Judy. "Not for the Faint at Heart: Contemplations on Usenet." Wired Women: Gender and New Realities in Cyberspace. Ed. Lynn Cherny and Elizabeth Reba Weise. Seattle: Seal Press, 1996. 126-140.
- Aycock, Alan, and Norman Buchignani. "The E-Mail Murders: Reflections on 'Dead' Letters." Cybersociety: Computer-Mediated Communication and Community. Ed. Steven G. Jones. Thousand Oaks: Sage, 1995. 232-238.
- Baym, Nancy. "The Emergence of Community in Computer-Mediated Communication." Cybersociety: Computer-Mediated Communication and Community. Ed. Steven G. Jones. Thousand Oaks: Sage, 1995. 138-163.
- Benson, Thomas. "Rhetoric, Civility, and Community: Political Debate on Computer Bulletin Boards." Communication Quarterly 44 (Summer 1996): 359-378.
- Camp, L. Jean. "We are Geeks, and We are Not Guys: The Systems Mailing List." Wired Women: Gender and New Realities in Cyberspace. Ed. Lynn Cherny and Elizabeth Reba Weise. Seattle: Seal Press, 1996. 114-125.
- Carstarphen, Meta G., and Jacqueline Johnson Lambiase. "Domination and Democracy In Cyberspace: Reports from the Majority Media and Ethnic/Gender Margins." Cyberghetto or Cybertopia? Race, Class, and Gender on the Internet. Ed. Bosah Ebo. Westport, Conn.: Praeger, 1998. 121-136.
- Clerc, Susan. "Estrogen Brigades and 'Big Tits' Threads: Media Fandom Online and Off." Wired Women: Gender and New Realities in Cyberspace. Ed. Lynn Cherny and Elizabeth Reba Weise. Seattle: Seal Press, 1996. 73-87.
- Grossman, Wendy. Net Wars. New York: University Press, 1997.
- Gurak, Laura. Persuasion and Privacy in Cyberspace: The Online Protests Over Lotus Marketplace and the Clipper Chip. New Haven: Yale University Press, 1997.
- Horn, Stacy. Cyberville: Clicks, Culture, and the Creation of an Online Town. New York: Warner Books, 1998.
- Ito, Mizuko. "Virtually Embodied: The Reality of Fantasy in a Multi-User Dungeon." Internet Culture. Ed. David Porter.

- New York: Routledge, 1997. 87-110.
- Kendall, Lori. "MUDder? I Hardly Know 'Er! Adventures of a Feminist MUDder." Wired Women: Gender and New Realities in Cyberspace. Ed. Lynn Cherny and Elizabeth Reba Weise. Seattle: Seal Press, 1996. 242-264.
- McGee, Jennifer J. Net of A Million Lies: Rhetoric and Community on Three Usenet Newsgroups. Unpublished Ph.D. dissertation, University of Minnesota, 1997.
- McLaughlin, Margaret L., Kerry K. Osborne, and Christine B. Smith. "Standards of Conduct on Usenet." Cybersociety: Computer-Mediated Communication and Community. Ed. Steven G. Jones. Thousand Oaks: Sage, 1995. 90-111.
- McRae, Shannon. "Coming Apart at the Seams: Sex, Text, and the Virtual Body." Wired Women: Gender and New Realities in Cyberspace. Ed. Lynn Cherny and Elizabeth Reba Weise. Seattle: Seal Press, 1996. 242-266.
- Mitra, Ananda. "Diasporic Web Sites: Ingroup and Outgroup Discourse." Critical Studies in Mass Communication 14: 1997, 158-181.
- Mitra, Ananda. "Virtual Commonality: Looking for India on the Internet." Virtual Culture: Identity and Communication in Cybersociety. Ed. Steven G. Jones. Thousand Oaks: Sage, 1997. 55-79.
- Odzer, Cleo. Virtual Spaces: Sex and the Cyber Citizen. New York: Berkley Books, 1997.
- Reid, Elizabeth. "Communication and Community on Internet Relay Chat: Constructing Communities." High Noon on the Electronic Frontier: Conceptual Issues in Cyberspace. Ed. Peter Ludlow. Cambridge: MIT Press, 1996. 397-411.
- Reid, Elizabeth. "Virtual Worlds: Culture and Imagination." Cybersociety: Computer-Mediated Communication and Community. Ed. Steven G. Jones. Thousand Oaks: Sage, 1995. 164-183.
- Schmitz, Joseph. "Structural Relations, Electronic Media, and Social Change: The Public Electronic Network and the Homeless." Virtual Culture: Identity and Communication in Cybersociety. Ed. Steven G. Jones. Thousand Oaks: Sage, 1997. 80-101.
- Shaw, David. "Gay Men and Computer Communication: A Discourse of Sex and Identity in Cyberspace." Virtual Culture: Identity and Communication in Cybersociety. Ed. Steven G. Jones. Thousand Oaks: Sage, 1997. 133-145.
- Tepper, Michele. "Usenet Communities and the Global Politics of Information.@" Internet Culture. Ed. David Porter. New York: Routledge, 1997. 39-54.
- Turkle, Sherry. Life on the Screen: Identity in the Age of the Internet. New York: Simon and Schuster, 1995.
- Watson, Nessim. "Why We Argue About Virtual Community: A Case Study of the Phish.Net Fan Community." Virtual Culture: Identity and Communication in Cybersociety. Ed. Steven G. Jones. Thousand Oaks: Sage, 1997. 102-132.
- Zickmund, Susan. "Approaching the Radical Other: The Discursive Culture of Cyberhate." Virtual Culture: Identity and Communication in Cybersociety. Ed. Steven G. Jones. Thousand Oaks: Sage, 1997. 185-205.