Against the Grain

Volume 22 | Issue 3 Article 12

June 2010

Op Ed -- Trapped in the Web

Fred Kameny
Duke University Press, fkameny@dukepress.edu

Op Ed — Trapped in the Web

by **Fred Kameny** (Duke University Press) <fkameny@dukeupress.edu>

Not long ago, as I was reading **Max Blumenthal's** *Republican Gomorrah* (2009), I came across the following endnote:

> **Laurie Goodstein**, "Issuing Rebuke, Judge Rejects Teaching of Intelligent Design," *New York Times*, December 21, 2005, *http://query.nytimes.com.gst/fullpage.html?res=9C00E6DC1430F932A15751C1A9639C8B63&sec=&spon=&pagewanted=all*. Also, for full **Jones** opinion, see *http://msnbcmedia.msn.com/i/msnbc/sections/news/051220_kitzmiller_342.pdf*.

What a mess. Yet the author was merely citing an article in *The New York Times* and an opinion by a federal judge. The endnote would have been shorter, more readable, and more informative if it had looked like this:

> **Laurie Goodstein**, "Issuing Rebuke, Judge Rejects Teaching of Intelligent Design," *New York Times*, December 21, 2005, § A, p. 1; for full **Jones** opinion see ***Kitzmiller v. Dover***, 400 F. Supp. 2d 707 (M.D. Pa. 2005).

Instead I was assaulted by a mass of alphanumeric garble. The reason is that the author was doing what everyone else seems to be doing these days: when documenting a source to which access has been gained by electronic means, focusing on the means rather than the source.

I submit that this practice is ill-advised, and that in general authors and their editors spend too much time worrying about the fine points of digital citation. I am admittedly biased toward print, having been in the business of publishing books for more than a quarter-century. And there is no question that digital sources are multiplying, that many scholars are infatuated with them, and that by some accounts the printed book itself may be headed for extinction. Yet even electronic sources can be cited in ways that are more or less helpful to the reader. And if we do not rein in the current practice of reproducing long strings of code as if we were all would-be computer programmers, notes and bibliographies will become increasingly indecipherable and useless.

In discussing how to handle digital sources I propose that we begin by distinguishing between those that have printed analogues and digital sources that exist in digital form alone.

If a digital source has a printed analogue, then I think the course is clear: cite the printed source, not the digital one. In an important sense the digital counterpart to the printed source does not exist — it is simply an electronic epiphenomenon. Citing to the Web edition of a major newspaper like *The New York Times* or the *Washington Post*, as the author quoted above has done, seems especially pointless, since several hundred thousand copies of the print edition are produced daily. (I would make that point even more strongly when the Web edition of a publication is a PDF of the print edition, as is often true of government documents, and as is increasingly so of newspapers.) Let us put aside for the moment the question of how to handle material that is posted on the Website of a newspaper but never appears in the printed edition — more about that presently.

Now, one might say that intellectual honesty demands citing to the electronic edition if that is the edition one has consulted. My response is that the author should consult the printed edition in any event. But what if the issue of the newspaper being cited is more than a few months old? In that case finding hard copy will be almost impossible, and one will need to resort to microfilm. Here we have a solution to the puzzle. For decades scholars have been citing microform editions of newspapers, for the most part without making any mention of the medium. I think that omission is defensible because specifying the medium is generally not relevant to the reader: I don't need to know whether the author read *The New York Times* on paper or on microfilm or on the Web, just as I don't need to know whether the author was wearing contact lenses while doing his or her research (and yet, it is plausible that failing to wear contacts is more likely to result in mistaken transcription than deciding to read the *Times* on microfilm rather than in hard-copy form). In other words, the chances of there being a material difference between print and microform are vanishingly small; and the chances of there being a material difference between print and Web are only slightly greater. I would not, incidentally, make the same argument if the work being cited were a microform edition of the *First Folio*, or *Early American Imprints*: in the first instance, textual exactitude is inherently important; and in the second, consulting the microform edition is pretty nearly the only way to gain access to the contents of the series.

In case anyone thinks I am being too cavalier about citing printed sources even if one has not laid eyes on them, let me draw on an analogy from the legal world. As many readers will know, lawyers are fastidious about citation, and even seemingly trivial lapses in citation have been known to get people in hot water with the professor, the boss, and even the judge. As readers probably also know, nearly all legal research is now done online, through **Lexis** and **Westlaw**, which are the duopolists of the legal database trade. If a lawyer needs to cite a U.S. Supreme Court case, he or she will pull up the case on a computer screen rather than go to the library to find the printed edition. That makes perfect sense. The online edition is essentially the same as the printed edition, even showing where each page of the printed edition begins. Although the printed edition is still published and probably always will be, I honestly don't know who uses it anymore. But here is the interesting thing. If you look up ***Brown v. Board of Education*** in an online legal reporter, you are not expected to specify in your brief, or your law-review article, that you have used the **Lexis** edition of ***Brown v. Board of Education***, and that you accessed it on 28 June 2008: instead you are expected to cite **Brown** in exactly the same fashion as if you had consulted a hard-copy edition — in other words, you are expected to do precisely what I am suggesting we do when citing scholarly material in general. Indeed, if you add extraneous information such as the reference to **Lexis** and the access date, you are, strictly speaking, citing the case incorrectly. It is always possible that there will be discrepancies between the **Lexis** or **Westlaw** edition and the printed edition — I have seen minor ones. But even so, the legal profession has decided that it can live with that small degree of imprecision and that the citation of legal materials should be, in the words of the **American Association of Law Libraries**, "vendor and media neutral." And so academics, and those of us who edit their work, need to ask ourselves whether it is really necessary for us to be more pedantic than the legal profession.

This brings us to the citation of sources that exist only in electronic form. Because of my bias in favor of print, I always advise authors who bother to ask that they should print out any Web pages they have consulted, and then, whenever they cite those pages, append the notation "printouts on file with author." That much is uncontroversial. But many of my publishing colleagues take issue with what I see as a logical corollary to that recommendation: that transforming

continued on page 43

Op Ed
from page 42

a Web citation into a print citation obviates citing the URL, and that a generic citation like one of the following will suffice:

Website of the World Health Organization

Website of the World Health Organization (visited 1 May 2007)

Website of the World Health Organization (printouts on file with author)

Website of the World Health Organization (visited 1 May 2007; printouts on file with author)

I even regard the access date as optional, because just as I am not really concerned whether the author has read *The New York Times* on paper or on the Web, I am not concerned whether the author has visited a given Website on 1 May or 10 May, especially if there is a hard-copy record of the site. Again, the counter-argument is that citing the URL and the access date is a matter of scholarly thoroughness; and again, my counter-counter-argument is that thoroughness always has its limits. When we cite a widely available printed book or journal, we do not cite the repository where we used it; neither do we indicate which printing of a given edition of a book we have consulted, even though there can be important differences between printings. Similarly, when we cite a telephone interview, we give the name of the person interviewed and the date of the phone conversation; we do not give the telephone number at which the person was reached or the time of day when the call was made, nor do we specify whether the telephone was "corded," cordless, or cellular (and yet, as with the contact lenses, that distinction could be relevant to how accurately the interview was transcribed).

Well, you might ask, what about Websites that are constantly changing? A colleague of mine mentioned the example of a newspaper's Website on election night: the site changes every few minutes, whenever there is a significant number of returns from previously unreported precincts. In my view, examples like this only strengthen the case for printing out Web pages. They also demonstrate that once we start down the road of exhaustive electronic citation, we will never get to the end of it: if the newspaper's Website changed thirty or forty times on election night, it does little good to tell the reader that you accessed the site on 5 November: you will also need to specify at what time you accessed the site — of course remembering to specify as well whether you mean Eastern or Central or Mountain or Pacific time, or something else.

One reason why I disapprove so strongly of hyper-correctness in electronic citation is that the standards for print citation (and other kinds of documentation) have been getting so lax. If a hardcover book was published in 1960, I know many scholars who will think nothing of citing the paperback edition published in 1962 without even mentioning the original hardcover, simply because the paperback edition is the one that they happen to have in their office — a bibliographic lapse that to me is completely unacceptable. And for some really questionable advice we need look no further than the *Chicago Manual of Style*: among the practices that it condones are quoting from foreign-language sources only in translation, as if **Plato** has written in English; omitting the initial article in names of works to "fit the surrounding syntax" (resulting in references to **Joyce's** *Dead*, **Faulkner's** *Hamlet* and other absurdities); and silently changing the capitalization of the initial letter of a quotation. I find it incomprehensible that authors and publishers are willing to countenance this sort of editorial high-handedness, while at the same time insisting that anyone who cites a Web source without including the URL is a bad scholar.

I have made these arguments before, and one answer I often get is that the problem I am battling will soon take care of itself, because digital object identifiers (DOIs) and Websites like **tinyurl.com** and **Webcitation.org** will make serpentine citations like the ones at the beginning of this piece a thing of the past. Websites will be archived, so that readers will no longer need to worry about dead links; and identifiers will become shorter and easier to use. That seems like a happy prospect. But if one believes as I do that it is a shoddy practice to cite Web editions of readily accessible print documents, and that Web sources in general are being cited too indiscriminately, having more usable Web citations might make our problems worse, not better. So I continue to advocate three basic principles: First, cite the source, not the medium through which access to the source was obtained. Second, if citing the medium is unavoidable, choose a stable medium rather than an unstable medium — it is perverse to do otherwise. Third, if the source exists only in an unstable medium, translate it into a stable medium and then cite it.

There may be an added benefit. Cut-and-paste citation has contributed to cut-and-paste scholarship of a broader sort. More and more I am seeing manuscripts that cite sources like **Wikipedia** — the scholar's equivalent of the journalist's "Some say . . ." Perhaps being more rigorous in how we cite will make us more discerning in what we cite.

Some years ago **Judge Frank Easterbrook** of the U.S. Court of Appeals published an article called "Cyberspace and the Law of the Horse." This was at a time when the Internet was in its infancy and technological utopianism was rampant. We were assured that the Internet was going to change everything. It may be recalled, for example, that the old ways of valuating businesses were said to be obsolete — this time things were different. Even our legal system, several centuries in the making, would need to be completely reconfigured. **Judge Easterbrook's** article suggested that, at least in the legal realm, these speculations were nonsense. Legal rules would be applied to new sets of facts and circumstances, but the rules themselves would not change: just as there was no law of the horse, there was no law of the Internet. Eventually this prediction was borne out, as the legal system accommodated the technological advances of the late twentieth century without needing to be discarded wholesale. Yet academic writers and editors still seem to be in thrall to the dot.com propaganda of the 1990s: in this view we are on the frontier of changes so momentous, so unprecedented in scale, that all our assumptions and conventions will need to be overturned.

This way of thinking seems to me deeply misguided. As long as writers have been citing sources, the process has adhered to some fundamental precepts, all premised on the writer's obligation of good faith toward the reader: the citation must be correct, unambiguous, informative, and concise, and the reader must be able to retrace the author's steps, or at least be able to understand why doing so is impossible. Nothing about the digital world alters these simple truths. To believe otherwise is to lose sight of some of the basic principles of scholarship.

Rumors
from page 16

And, I understand that **Gail's** daughter, **Dr. Sandra Hirsch**, is the new **Director of the San José School of Library and Information Science**. As we know, **Sandy** is a second generation librarian whose library experience dates back more than 25 years, when she worked as a library assistant in an academic library. Later, she worked as a librarian in academic and special libraries, including a law library and a corporate library. More recently, she chaired the **Palo Alto (California) Library Advisory Commission**, which created a long-range library plan for the city that resulted in a voter-approved bond measure for improved libraries in Palo Alto, even in the midst of difficult economic times. **Sandy** holds a Ph.D. in Library and Information Science from the **University of California, Los Angeles (UCLA)** and a **Master of Information and Library Science (MILS**) from the **University of Michigan**. **Sandy's** appointment with the School will begin in August 2010. Current **SLIS Director Ken Haycock** will be retiring from **San José** at the end of this academic year, although he will continue working with other academic units on strategic planning and supervising doctoral students in the **San José Gateway Ph.D. Program**. I will never forget **Gail** telling me the story that she gave **Sandy** a

continued on page 49