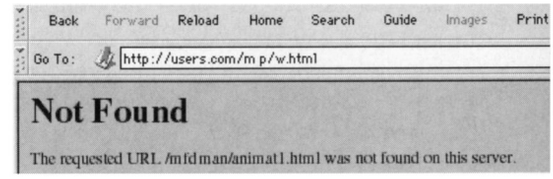
Books in Bytes? Not Yet
Author(s): Peter Givler
Source: *Academe*, Vol. 85, No. 5 (Sep. - Oct., 1999), pp. 61-64
Published by: American Association of University Professors
Stable URL: https://www.jstor.org/stable/40251772
Accessed: 29-03-2020 16:52 UTC

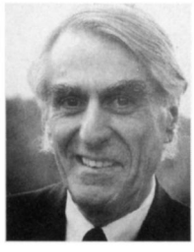
Books in Bytes? NOT YET



Electronic monographs may be the wave of the future. But publishers must first build an infrastructure for citations, copyrights, authentication, and preservation.

By Peter Givler

IN A MUCH-DISCUSSED ARTICLE PUBLISHED IN the *New York Review of Books* last spring, historian Robert Darnton proposed a new form of scholarly publication: a document that would exploit the unique attributes of electronic communications instead of merely imitating the flat, traditional world of print.[1] Darnton imagined a complex, interconnected work with different levels of information stacked one on top of the other to form a pyramid. The top level might report briefly on a historian's research and its findings, the next layer down could flesh out the arguments or present alternatives, and the third level would contain the documentation on which the work rests, along with interpretive essays. The fourth layer would feature discussions of theoretical and historiographical questions; the fifth, suggestions for teaching; and the sixth, an open-ended archive of reviews, editorial correspondence, letters from readers, and so forth.



Peter Givler is executive director of the Association of American University Presses.

Readers of such a work could explore this pyramid as fully as they wished. General readers might read only the top layer. Specialists, however, could examine alternative hypotheses, assessing the documentary evidence for themselves, or they could delve into the theoretical and historiographic issues that shaped the author's interpretation. The top layer might be made available as a paperback book, but any or all of the layers could be printed out as needed, and—perhaps most important—electronic links between layers could be developed and followed, and new meanings explored. Readers could skim across the surface of any level in the pyramid, but they could also burrow down through them—a possibility that, in Darnton's words, "would elicit a new kind of reading."

A new kind of reading? Darnton doesn't elaborate, but I understand him to mean that his model will begin to exploit one of the most fascinating possibili-

ties of electronic communication: its ability to present a web of information through which the reader, in theory, can navigate as effortlessly as thought arcing from one idea to the next. Darnton's model suggests an electronic artifact—one no longer constrained by the imperative of a linear, printed text to begin just here and end precisely there—that expands dramatically the mental space in which we read a book. At the same time, his model does not catapult its readers into a chaos of unmediated information but imposes a boundary on itself capable of giving a work shape and form. The work is still about this, and not that. It still starts here, and may even still end over there, although in granting its readers the freedom to follow a more complex and many-branched route from here to there, it allows new meanings and endings to emerge.

For all the demands this complexity may place on the reader, it has, for Darnton, a richly compensating virtue: it imitates the work of doing history itself. He writes:

> Any historian who has done long stints of research knows the frustration over his or her inability to communicate the fathomlessness of the archives and the bottomlessness of the past. If only my reader could have a look inside this box, you say to yourself, at all the letters in it, not just the lines from the letters I am quoting. If only I could follow that trail in my text just as I pursued it through the dossiers, when I felt free to take detours leading away from my main subject. If only I could show how themes crisscross outside my narrative and extend far beyond the boundaries of my book. Not that books should be exempt from the imperative of trimming a narrative down to a graceful shape. But instead of using an argument to close a case, they could open up new ways of making sense of the evidence, new possibilities of making available the raw materials embedded in the story, a new consciousness of the complexities involved in construing the past.

The initial costs of making the transition from print to electronic publishing will be high.

This is wonderfully heady stuff, and no doubt other scholars in the humanities and social sciences can imagine variations of Darnton's model that would open up creative possibilities in their own fields. University presses are eager to develop electronic technology to improve scholarly communication—to make it richer in meaning, better able to disseminate the results of scholarly research, more efficient, and less costly. In fact, we already know how to build the kind of monograph Darnton has proposed. We may lack a few of the tools we need to actually construct it, but developing them shouldn't be a major task.

So what's stopping us? Why aren't all scholarly publishers charging full speed ahead into the future that Darnton imagines? Part of the reason is economic. I am not going to parse the costs of electronic versus print publishing here, but I will heartily second Darnton's recognition that the initial costs of making the transition from print to electronic publishing will be high. The real reason we are cautious, though, is more fundamental. True, if we build it, people will probably come. But when they get there, what will they find?

Scholarly Traditions

KNOWING HOW TO BUILD A NEW ELECTRONIC WORK OF scholarship is one thing; having the infrastructure to support it after it's built is something else again. In the world of printed books and journals, an infrastructure of scholarly conventions, developed over centuries, already exists. It works so well and reliably that we barely think to question it. Consider that basic signifier of academic writing, the lowly footnote, in its most mundane form, the bibliographic citation. As anyone who publishes a scholarly book or article knows, the task of writing such a citation is not to be undertaken in an antic or rebellious mood. Whether the title of a cited article should be in italics or quotation marks, how the names are to be listed in a work of multiple authorship, what style of punctuation is to be used in a footnote and how it differs from the style to be used in a bibliography—all these and many more such details are specified in rules as niggling as they are absolute, and as arbitrary as they may seem inconsequential.

Niggling, certainly, and arbitrary, perhaps, but inconsequential? The rules impose consistency, and consistency ensures that the threads connecting a scholar's work to her sources and to the work of other scholars can be followed easily. Such ease of tracking allows a scholar's work to be evaluated and built upon. In this sense, scholarship—even though it may be best pursued in the solitude of a carrel or study—is always collaborative. It is a quest for the meaning of one's sources (and sometimes a search for those sources themselves), as well as a conversation with one's colleagues, living and dead, who have wrestled with similar questions. Bibliographic references provide a stable, clear, and public system of referencing other texts, and thereby of acknowledging one's sources and the work of one's peers. Bibliographic citations are the humble pegs that hold together the elaborate edifice of written scholarly communication. Without them, modern scholarship would be, quite literally, unthinkable. Unfortunately, a comparable infrastructure for the world of electronic scholarship is still under construction. We can't yet cross-reference electronic texts reliably. A system for doing so will come, but the difficulties are more basic than agreeing on a new footnote form for electronic documents.

First of all, though, why bother? Don't hyperlinks do for electronic documents what footnotes did for print, and do it

even better? Why simply point at a text when we can provide the means for readers to examine it for themselves?

Link Rot and Copyright Law

TWO DIFFICULTIES WITH HYPERLINKS DISQUALIFY THEM AS substitutes for bibliographic citations. The first is "link rot," or what happens to links if they're left unattended. A link works only when it connects you to something. Links break for whatever reason—loss of interest, a move to another site. A broken link tells you only that it's broken: "URL Not Found" or "No Such Site on This Server." Bibliographic citations, clank as they may, at least name texts clearly and unambiguously. They tell you what to look for. A rotted link tells you nothing. It just drops you off in a vacant lot.

The legal status of hyperlinking is also uncertain. Under the fair-use provision of current copyright law, scholars have generally been allowed, for purposes of illustration, explication, or argument, to incorporate into their own works small amounts of material held under copyright by others without seeking the permission of the copyright holder. But what happens when a link in one document takes the reader to the entire text of a separate, copyrighted document? Links strive to be seamless: click on a button or a section of highlighted text and the linked material should appear on your screen. Is this material incorporated into the document you're reading? If it is, that incorporation would appear to be far outside the bounds of anything that has previously been considered fair use. Or is the link simply a convenience that saves readers a trip to the library? As a matter of copyright law, the answer is not clear.

Electronic storage media—at least the ones we've invented so far—have the longevity of a fruit fly.

Hyperlink problems aside, the fundamental impediment to developing an electronic standard for bibliographic references is the oddly Borgesian difficulty of authenticating electronic texts. In the world of ink and paper, much scholarly labor has been spent to establish the degree of authority of written texts. The bedrock on which all other textual claims rest is the author's finished manuscript (or, if that no longer exists, the printer's proofs corrected in the author's hand, or some other version of the text that can be argued to best represent the words the author intended to appear on the page).

But in the e-world there is no intrinsic marker to distinguish an original electronic file from its copy, or even a way to distinguish a copy of a copy from a copy of an original. This would seem to make the distinction between "copy" and "original" irrelevant—if only electronic files were tamper-proof. They aren't, though, and the Internet has spawned an entire subculture of hackers apparently dedicated to proving they never will be. So, how can you tell that the particular document you're looking at is a faithful replica of the original—that it hasn't been accidentally, or even maliciously, altered? If you wanted to compare your "copy" with the "original," would you know where to find the original—and once you found it, how could you be sure it was, in fact, the original?

There are various ways to "mark" an electronic file so that its authenticity can be verified, but as yet no agreed-on standards for doing so. And little use is being made of the methods that do exist. Establishing authority is further complicated by what many see as a virtue of electronic communication: the potential it offers for virtually instantaneous comment, review, and correction. If several versions of a work are available because its author took advantage of readers' comments to revise it, which one is the authorized version? Are any of them, or do they just represent different stages in the evolution of the author's ideas?

Disappearing Data

THE ISSUE OF ESTABLISHING AUTHENTICITY IS CLOSELY related to the problem of preserving electronic texts. Paper, for all its apparent fragility, has proved to be astonishingly durable; Shakespeare scholars today can work with editions of his plays printed four hundred years ago that, with reasonable care (and absent the rude shocks of history), will still exist four hundred years from now. By contrast, electronic storage media—at least the ones we've invented so far—have the longevity of a fruit fly. Compact disks, the most common electronic storage medium now in use, have an estimated life span of twenty-five to forty years; to be safe, archivists figure they're good for about ten. When the disk goes, everything encoded on it becomes instantly unreadable, lost beyond the power of any scholarly or electronic tool to recover it, as dead as disco. We have as yet no plastic counterpart for archival paper, and no amount of special care will prolong the life of a compact disk in the way that controlling temperature and humidity and light will prolong the life of a book.

What to do? The generally accepted solution is to "migrate the database": to copy everything one wants to preserve onto new generations of storage media as the old forms age toward unreliability. In theory, it's simple, even elegant. Unlike print, which is indissolubly wedded to the page on which it appears, an electronic file can be lifted intact from one storage medium and set down, perfect and complete, on another. We no longer need to preserve texts by prolonging the life of the artifact on which they have been inscribed. No more back-up microfilming, no more de-acidification programs, no more environmentally controlled stacks with limited access. All we need is one gizmo to read from the old medium, another to write to the new, and a computer in between to manage the process.

Absence of Political Will

THAT'S THE THEORY, AND WE HAVE THE basic technology to begin applying it. Nevertheless, little is being done. A few electronic archiving projects are under way, but they will preserve only a trickle in the sea of information that surrounds us. The obstacle appears to be one of political will: as a society we know what we need to do, but we haven't yet made the commitment to do it.

That said, if the will were there tomorrow, the difficulties in implementing it would not be trivial. Archiving electronically will require significant resources. As straightforward as the theory sounds, it will still demand a continuing investment in hardware, software, and trained personnel. In addition, for the foreseeable future, the new cost of electronic archiving will have to be paid on top of the cost of maintaining existing print archives. Finally, the traditional roles of publishers and librarians may have to change to allow the systematic archiving of electronic documents.

Until we resolve these difficulties, we still have books. At the very least, they are our cultural back-up files in a volatile era of transition. It would be an unspeakable tragedy if we were to leave the world of print before we could develop the necessary electronic analogues for the stability, ease of authentication, and forms of connection it offers. Without such analogues, we run the danger of losing not just scholarship, but also our scholarly selves.

Darnton's blueprint offers an exciting glimpse of the future, and we will learn a great deal from the pilot program to realize it that the Andrew W. Mellon Foundation recently funded. As we move forward, though, we need to devote the same energy and resources to building an electronic scholarly infrastructure that we do to creating new forms. Only when we have those vital underpinnings will we be able to fully explore the capabilities of the new media.

What will scholarship be like in the electronic future? It will embody new relationships between author and reader, new ways of looking at sources, new ways of constructing arguments, new ways of reading and even, perhaps, of doing scholarship itself. And it will sustain us with the homely and nourishing virtues of plain, old-fashioned books.

Note

1. Robert Darnton, "The New Age of the Book," *New York Review of Books,* 18 March 1999, 5–7.