School of Library and Information Science Faculty
Research Publications

School of Library and Information Science

9-1-1993

Libraries and Texts in the Electronic Environment

Gordon B. Neavill

Wayne State University, neavill@wayne.edu

Scholarly Communication

in an

Electronic Environment

Issues for Research Libraries

Robert Sidney Martin
Editor

Libraries and Texts
in the Electronic Environment

Gordon B. Neavill

It is generally recognized that we are in the early stages of a communications revolution comparable to the invention of writing, the transition from the roll of the ancient world to the medieval codex, and the invention of printing from movable type. Electronic texts in digital form are already an important part of our information environment, and they are likely to be an increasingly important part of it in the decades to come. It is too early to predict whether digitally encoded texts will eventually supplant traditional formats or coexist harmoniously with them. We know enough about the consequences of earlier communications revolutions, however, to know that the digital revolution will transform scholarly and intellectual life in ways we cannot begin to comprehend. Much will be gained, much that we now value will be lost; nothing will remain the same.

Jay David Bolter has explored the characteristics of electronic texts in his recent book, *Writing Space: The Computer, Hypertext, and the History of Writing*.[1] Bolter argues that electronic texts are not simply traditional texts in digital form but a completely new kind

of text. I'd like to begin by quoting several of his statements about electronic texts.

> A true electronic text . . . is not a fixed sequence of letters, but is instead from the writer's point of view a network of verbal elements and from the reader's point of view a texture of possible readings. (p. 5)

> The conceptual space of a printed book is one in which writing is stable, monumental, and controlled exclusively by the author. . . . The conceptual space of electronic writing, on the other hand, is characterized by fluidity and an interactive relationship between writer and reader. (p. 11)

> The vanishing of the fixed text alters the nature of an audience's shared experience in reading. All the readers of *Bleak House* could talk about the novel on the assumption that they had all read the same words. No two readers of an electronic book can make that assumption; they can only assume that they have traveled in the same textual network. (p. 8)

> Because an electronic text is not a physical artifact, there is no reason to give it the same conceptual unity as a printed book, no reason not to include disparate materials in one electronic network. (p. 7)

True electronic writing, in Bolter's view, "is not limited to verbal text: the writable elements may be words, images, sounds, or even actions that the computer is directed to perform." (p. 26) That is, the true electronic text is hypertext, which Bolter defines as "the interactive interconnection of a set of symbolic elements. . . ." (p. 27) The electronic text is not linear but layered; it is not a road to a predetermined destination but a network of possibilities.

As a classicist, Bolter is well aware of the similarities between electronic texts as he describes them and Greek oral poetry. He writes:

> The contrast between oral and written texts is important for an understanding of electronic writing, because in some ways the new medium more closely resembles oral discourse than it does conventional printing or handwriting. (pp. 58-59)

And finally:

> Electronic text is the first text in which the elements of meaning, of structure, and of visual display are fundamentally unstable. . . . All information, all data, in the computer world is a kind of controlled movement, and so the natural inclination of computer writing is to change, to grow, and finally to disappear. (p. 31)

Electronic texts of this kind make possible new literary genres founded on interactivity between readers and texts, and they open up new and potentially exciting modes of human communication. They also raise disturbing questions about what society's stock of knowledge will be like in the electronic environment.

Will stable, linear texts whose content is not in flux exist in the electronic environment? Will the distinction between authoritative and corrupt texts begin to fade, to be replaced by a recognition that a work can exist in endless variants, each of which may be equally valid? Will authorship be increasingly regarded as a collective activity as readers become accustomed to adapting or adding to the content of preexisting texts?

What will historical scholarship be like in the electronic environment? In the paper-based environment historical scholarship is based on documentary records that have survived from the past, records whose existence as physical objects continues long after the purposes for which the records were created have been served.

Many of these records have survived for extended periods unused or unknown before their research value was recognized and they were acquired by libraries or archives. Will documentary records of this kind exist in the electronic environment, or will we have to rely on oral tradition or its electronic equivalent to link us with the past?

The technological imperatives of the electronic environment favor revision, updating, and continual currency, not retrospective documentation. Magnetic media are easily erased and reused. Digital encoding of information carries this a step further, allowing the erasure of specific data within a record and the interpolation of new data at will. Outdated information in continually updated electronic databases such as city directories, membership lists, commercial catalogs, bibliographies of books in print, and telephone directories is likely to be erased. In the paper-based environment back volumes of these publications provide important historical documentation to scholars. In the electronic environment these databases become electronic palimpsests.

Before pursuing questions about the stock of knowledge in the electronic environment, it may be useful to reexamine how society has managed its stock of knowledge in the paper-based environment. The library is the primary institution by means of which literate societies retain contact with and control over their accumulated stock of recorded knowledge. Libraries acquire, organize, and provide access to the physical objects—books, manuscripts, etc.—in which knowledge is recorded. We try to preserve as best we can those objects whose content seems to be of permanent value. Over time a sizable stock of recorded knowledge is accumulated, and it is this stock of recorded knowledge—that which has been acquired and preserved in libraries—that in large part constitutes the stock of recorded knowledge that a given society can be said to possess.

The librarians whose day-to-day activities in the paper-based environment contribute so much to the shaping and codification of society's stock of recorded knowledge are not overly conscious of

this mission. Their workaday concerns tend to be practical and mundane and focused on serving the current information needs of library users and, to a much lesser extent, the anticipated needs of future users. The materials they acquire have two characteristics in common: in most cases they are readily available through established trade channels, and they are physical objects such as books and sound recordings that lend themselves to storage and preservation in libraries. CD-ROM's have been added to the arsenal of formats in recent years, and libraries are also increasingly utilizing database vendors to make information available that the library itself does not own. Rare book and manuscript librarians are unique in the extent to which they acquire materials outside commercial channels, by cultivating private collectors who may donate ready-made collections to the library, and by establishing contacts that may lead to the acquisition of unpublished archival materials. All this pretty much defines the scope of the stock of knowledge available in libraries. It is uncommon for libraries to assume responsibility for acquiring and preserving information that is not prerecorded and conveniently packaged, though a small number of activist institutions have done just that.

Institutions that have taken the initiative to capture and preserve unrecorded information include the Television News Archive at Vanderbilt University and the Archive of Folk Song at the Library of Congress with its program of field recordings. Because of these activities evening news programs and a wealth of folk music can be counted as part of our stock of recorded knowledge.

Things will be more complicated in the electronic environment. We as librarians will have to approach our role of shaping and preserving the stock of knowledge in a far more conscious and deliberate way than we do now, and our activities are likely to be closer to the activist model exemplified by the Vanderbilt Television News Archive and the Archive of Folk Song. If electronic texts are fundamentally unstable and have a natural inclination to change,

grow, and finally to disappear, we will have to find ways to capture and stabilize them if future generations are to inherit documentation from the past. We will have to make difficult decisions about which texts to transmit to the future and which texts future generations can do without, and we will have to make these decisions quickly before the texts evolve or disappear.

None of this will be easy. The whole concept of preservation seems to be antithetical to the electronic ethos. Attempts to capture and preserve electronic texts are likely to place librarians in the kind of conflict with vendors that the Television News Archive experienced with CBS Evening News. Some of the scholars for whom electronic texts are being preserved may be unsympathetic. The integrity of texts is of less concern in contemporary critical theory than the ways in which readers respond to them. Some scholars may find living, evolving texts more interesting than texts that have been captured and preserved like butterflies pinned in a case. We may have to struggle against currents of scholarly fashion as well as the technological imperatives of the electronic environment in order to transmit a usable stock of knowledge to future generations. It is well to bear in mind that fashions change and vary considerably from one discipline to another. If contemporary critical theory seems to devalue the integrity and authority of texts, we need only to turn to musicology to find the opposite trend. Here the emphasis is on analyzing original scores, chipping away encrusted performance practices and interpretations to recover as fully as possible the music as the composer originally intended it to be heard.

Many of the distinctions that we take for granted in the paper-based environment are likely to be blurred in the electronic environment. The distinction between recorded and unrecorded knowledge is less meaningful in an environment where a high percentage of communication other than face-to-face conversation is captured at least fleetingly in digital form and where electronic texts have much in common with oral tradition. The distinction between private and

public spheres of knowledge is less clearcut in the electronic environment. Are electronic lists and bulletin boards examples of private communication or do they constitute a public record that should be preserved? When the distinction between private and public knowledge is blurred, publication—which can be defined as the formal process by which works are made public, that is, transferred from the private possession of the author into the public stock of knowledge—becomes blurred as well. Many messages that remain unpublished in the traditional sense will increasingly be made available electronically to small or large groups. Our actions in shaping and preserving the stock of recorded knowledge in the paper-based environment have been based on relatively unarticulated distinctions between private and public knowledge, recorded and unrecorded knowledge, published and unpublished knowledge. Deciding what to preserve and what to ignore in the electronic environment will be far more difficult.

One option, of course, is for libraries to decline responsibility for the acquisition and preservation of electronic texts and to let them fend for themselves—to treat them, in other words, as libraries in the paper-based environment have treated orally transmitted knowledge. Oral tradition, it must be noted, has managed quite well without assistance from libraries and librarians.

A significant part of the stock of knowledge of any society, even the most literate societies, consists of unrecorded knowledge that is transmitted orally. Oral tradition can be a risky way of transmitting knowledge; any break in the chain of transmission from one generation to another results in an irretrievable loss of knowledge. And except for some orally transmitted sacred texts, the concept of authenticity is alien to oral tradition: each oral rendering of a given work represents a variant that can be regarded as a text in its own right. Yet oral tradition can be an extraordinarily robust form of transmission and preservation. Orally transmitted works have travelled from one end of the earth to the other and have survived

for centuries. To some extent, of course, oral transmission has been reinforced and extended in the twentieth century by radio, television, and sound recording—media that Walter Ong characterizes as "secondary orality."[2]

The stock of orally transmitted knowledge can be divided into two categories. First is the body of practical knowledge, attitudes, values, and the like that are passed on largely by example and word of mouth rather than the written word. Most people don't learn to perform practical activities such as carpentry or bookbinding by reading about them. Achieving a high level of skill at such pursuits generally requires direct contact with persons who possess the skills being sought. This is true of scholarly and intellectual life as well. John Ziman has written, "A mature scientist takes decades of training, and is the heir of subtle intellectual traditions. Academic institutions are governed largely by unrecorded principles, handed on from father to son, from master to pupils, in the intimacies of the seminar room, the study and the laboratory."[3] Oral history projects occasionally try to capture this kind of orally transmitted knowledge, but for the most part such knowledge remains unrecorded.

The second category of oral tradition consists of oral texts, including sacred works, epic poetry, songs, jokes, and tales. These works fall under the broad definition of "text" that D. F. McKenzie advanced in his Panizzi Lectures. Basing his definition on the meaning of the Latin *texere*, "to weave," McKenzie regards any work woven of words or other symbols as a text, including "verbal, visual, oral, and numeric data, in the form of maps, prints, and music, of archives of recorded sound, of films, videos, and any computer-stored information."[4]

Oral texts of this kind remain very much alive in our society. A recent recording by Bobby Rush, a funky, contemporary blues singer based in Jackson, Mississippi, provides a good example. The artist is a popular figure on the "chitlin' circuit" of black clubs in Mississippi and neighboring states; he is not your stereotypical

"traditional" bluesman. The recording, called "Playin' Me Crazy,"
appears on an album issued in 1988. In the song, Bobby Rush
comes home "tired as a man can be" to find on three successive
nights strange shoes under his bed, another coat on the coat rack
where his coat should be, and an unfamiliar car in the driveway.
Each night he awakens his wife and asks her to "explain this stuff to
me." She responds that he's being silly and claims that what he sees
are two cabbage heads, a blanket, and a large tricycle. Bobby Rush
remains unconvinced; he's been all over the world "from Maine to
Mexico" and has never seen cabbage heads with shoelaces, a blanket
with sleeves, or a tricycle with four wheels.[5]

The song has been part of the blues tradition for more than sixty
years. The first recording of it, released under the title "Drunkard's
Special," was made in Dallas in 1929 by a singer named Coley
Jones. There have been a number of other recorded versions, the
best known probably being Sonny Boy Williamson's "Wake up
Baby" recorded in Chicago in 1958. As you would expect, Bobby
Rush's version is not identical to any of the others. It's been
updated: Bobby Rush comes home to find a car in his driveway
where his car should be; in other versions it's a mule in the stable.
He includes the "cabbage head" image which is missing from Sonny
Boy Williamson's version but is central to all the other versions I'm
familiar with—but he introduces it at the beginning, where it's
offered as an explanation of strange shoes under the bed. Most
versions culminate with the singer coming home and finding another
head on the pillow, which his wife protests is nothing but a cabbage
head. My purpose here is to point out that the song exists in a
number of variants and has been a living part of the American blues
tradition probably as long as that tradition has existed.

But the song is older than the blues, and it does not belong solely
to the black musical tradition. It's the common property of whites
and blacks; there are a number of versions of it by white folk
singers. But what's really interesting is that Bobby Rush's "Playin'

Me Crazy" is simply the latest variant of "Our Goodman," no. 274
in Francis James Child's *English and Scottish Popular Ballads*.[6]
Child cites versions of the song that were transcribed in Scotland in
the eighteenth century; he also notes variants in Gaelic, Flemish,
German, the Scandinavian languages, French, Italian, and Catalan.

It seems to me that this may have some relevance to the subject
of electronic texts. It is possible for certain kinds of texts to survive
for centuries completely without benefit of libraries. That is a
chastening thing for a librarian to admit! Of course, the only way
we know any of this is through the recorded versions—the printed
ballads that Child studied, and the sound recordings that document
the subsequent career of "Our Goodman" in the twentieth century.

I have said that the documents acquired and preserved in a
society's libraries in large part constitute the stock of recorded
knowledge that the society can be said to possess. When we think
about selecting the documents that will constitute the stock of
recorded knowledge, it is important to distinguish between scholarly
publications on the one hand and documents that constitute the raw
material for scholarly research on the other. Studies of scholarly
communication tend to focus on scholarly publications themselves,
but the real problems in the electronic environment concern the raw
material for scholarly research, especially for humanistic and
historical research. Despite the exponential growth in scholarly
publishing that has taken place since the modern research library
came into existence just over a century ago, scholarly publications
constitute a clearly identifiable and finite body of recorded knowl-
edge. The electronic environment poses potential hazards for
scholarly publications, but the academic community recognizes their
importance and is sufficiently powerful that, one way or another,
access to scholarly publications in the electronic environment is
likely to continue and may be enhanced.

The raw material for scholarly research is another matter. Here
we are not dealing with a finite or clearly identifiable body of

recorded knowledge, and the research value of some of this material is not immediately obvious.[7] The fact is that almost anything can be of research value to somebody. This is especially true at a time when scholarly horizons are expanding and subjects that used to be ignored or dismissed with contempt—popular culture, new kinds of social history, gender studies, multicultural studies, sports history and the like—are attracting widespread attention.   To further complicate matters, the raw material for scholarly research includes unpublished as well as published documents. There is no way that a society's research libraries can acquire and preserve everything of potential research value. Our collections in this area tend to be hit or miss affairs, distressingly weak in some areas and astonishingly rich in others, idiosyncratic, and often unique. We could always do better, but we make do with what we have. Ultimately the state of our scholarly knowledge, especially in the humanities, depends on how effectively we acquire and preserve the raw materials needed for humanistic and historical research.

Many of the most important collections of these primary sources now in our research libraries were formed in the first instance by private collectors. I am thinking not so much of collectors like Henry Huntington or Pierpont Morgan but of George Thomason, the London bookseller who systematically collected—and dated—more than 22,000 pamphlets and fugitive pieces that appeared between November 1640 and April 1661, and whose collection, which found its way into the British Museum nearly a hundred years after his death, is the single most important source for charting the political currents of that turbulent period.[8]   Another example is Samuel Pepys, whose collection of seventeenth-century ballads and chap-books—the "penny merriments"—formed a small but important part of the 3,000 volumes he willed to Magdalene College, Cambridge.[9] Closer to home, Arthur A. Schomburg's collection of books, manuscripts, and prints documenting the African-American experience, acquired initially in the early decades of the twentieth century,

became the nucleus of what is now the Schomburg Center for Research in Black Culture.[10] There have been thousands of private collectors like Thomason, Pepys and Schomburg. Without such collectors many movements, genres and events could not be subjected to substantive scholarly analysis.

Private collectors are often in a better position than institutional collectors to form pioneering collections like these. The institutional collector operates under inherent constraints. One is the constraint of responsibility: you're spending other people's money and you have to justify your actions to persons who may not be blessed with scholarly imagination. Another constraint is that you have to be responsive to a broad constituency: you can't ride hobby horses and devote all your resources to a single collection the way private collectors can. Furthermore, institutions tend to be conservative; private collectors often perceive the significance of emerging areas long before academics discover them. And finally, private collectors sometimes have more money to spend than institutional collectors. The librarian's unique role is to take such collections, organize them, preserve them, and make them accessible to users as far into the future as possible.

What will the private collector be like in the electronic environment? This a difficult subject to speculate about. In the paper-based environment, the impetus to collect is intertwined with the inherent appeal of the objects collected, the pleasure of the chase, and the satisfaction of acquiring objects that are rare or that collectively define or illuminate a subject area. In the electronic environment the distinction between an original and a copy is meaningless. We won't be collecting first editions. But there will no shortage of texts to pursue. Some texts will cease to exist, but many will simply be tucked away and forgotten. Some collectors of the future may resemble the manuscript hunters of the early Renaissance who prowled decayed monasteries searching for lost classical texts. The quest will be complicated by the limited life expectancies of digital

storage media and problems of software and hardware dependency. The collector may not know what an old tape or disk contains until its storage code is broken and the text is reformatted so that it can be accessed by current software and hardware.

George Thomason is probably a likelier model for the collector of electronic texts than the Renaissance book hunter Poggio Bracciolini. Collectors who recognize the importance of texts when they are current, bring them together, and preserve them in electronic databases will be the ones who document the social, political, literary, and intellectual movements of the twenty-first century. They are the ones who will make possible much of the humanistic scholarship of the future.

One type of collection that may diminish in importance is the working library of the individual scholar. In the paper-based environment institutional collections are commonly enriched by donations of scholars' working libraries. Here I am thinking about utilitarian libraries brought together to support a scholar's research and writing rather than collections consciously undertaken to document something. Working libraries of electronic texts will occupy little space, they won't be costly to move, and there will be less reason to give them to institutions when scholars conclude research projects, retire, or die.

Even if someone donated a working library of electronic texts to an institution, it would probably be of limited use. In the paper-based environment if a portion of a book is relevant to your needs you buy the whole book. You get a coherent package of recorded knowledge that different people can use in different ways. But in the electronic environment you download only what you need—a chapter or two, a page, or a single paragraph.   It would be exciting to get a library like this from Isaac Newton or Thomas Wolfe, but for general purposes a collection of electronic scraps would be useless. It would be worse than getting a large collection of miscellaneous photocopies or offprints.

I have argued that libraries have an institutional responsibility for our accumulated stock of recorded knowledge and that they are the only social institution with that responsibility.[11] I do not think this will change in the electronic environment. Commercial information vendors have never been responsible for preserving recorded knowledge. The fate of the Frederic G. Melcher library on the history of the book trade is a good example of this. Melcher left the collection, which John Tebbel described as "the finest private collection on the subject in the country. . . . it surpasses all the public collections as well, including even that of the Library of Congress,"[12] to the R. R. Bowker Company, of which he had been president. For many years it was open to scholars. The last time the Bowker Company changed hands its new owners moved the firm to smaller quarters and sold the Melcher library to a bookseller. Subsidizing a scholarly research collection was not high on the list of corporate objectives, and the Melcher library has since been dispersed.

Electronic information vendors, like print publishers, are in business to provide information that people will pay for. Print publishers do not keep books in print forever; electronic information vendors are not likely to retain texts that are no longer in demand. Information that is no longer current and nonscholarly texts that have served the purposes for which they were created will be especially vulnerable. Electronic information vendors share another characteristic with print publishers: many of them will go out of business. When that happens even current information may cease to be available. Unless we create new institutions—perhaps some sort of Center for Research Libraries specializing in electronic texts—research libraries cannot escape responsibility for the accumulated stock of recorded knowledge. This is especially true for the source materials that will be needed by future humanistic and historical scholars.

The acquisition and preservation of electronic texts in libraries will raise very difficult economic issues. As I have noted elsewhere, "In the print-based environment, current publications constitute the vast majority of acquisitions at most research libraries. Retrospective information needs are served in large part by materials whose original purchase was justified on the basis of their provision of current information. The economic link between the provision of current and retrospective information is broken in the electronic environment where libraries provide access to a wide range of computer-based records they do not own."[13] Acquisitions budgets in academic libraries are already beginning to shift from ownership to electronic access, especially in the area of serials. In the long run this trend may enhance access to current materials. But money spent on electronic access—renting instead of buying—yields no intellectual equity. It contributes nothing toward the creation of a stock of retrospective knowledge.

If research libraries are to retain responsibility for the stock of recorded knowledge in the electronic environment, increased economic support specifically for the acquisition and preservation of retrospective materials will be required. Retrospective materials in digital form will not have to be extensively duplicated and can be shared in a network environment. But electronic access is possible only to texts that somebody owns. It will not be easy to allocate limited resources to materials that will be used infrequently and mostly by future generations, but the money will have to be found if humanistic and historical research is to survive.

The transition to an electronic environment will be hardest for libraries whose responsibilities extend beyond the provision of current information. The technological challenges will be minor compared with the conceptual, organizational, and economic challenges. If the transition is accomplished, special collections departments may find themselves assuming primary responsibility for all kinds of retrospective materials. In that event they will be even

more central to the concept of the research library than they are today.

Notes

1. Jay David Bolter, *Writing Space: The Computer, Hypertext, and the History of Writing* (Hillsdale, N.J.: Lawrence Erlbaum Associates, 1991).

2. Walter J. Ong, *Orality and Literacy: The Technologizing of the Word* (London and New York: Methuen, 1982), 136.

3. John Ziman, *Public Knowledge: An Essay Concerning the Social Dimension of Science* (Cambridge: Cambridge University Press, 1968), 88.

4. D. F. McKenzie, *Bibliography and the Sociology of Texts* (London: British Library, 1986), 5.

5. Bobby Rush, "Playin' Me Crazy," *A Man Can Give It But He Can't Take It* (Jackson, Miss.: LaJam Records LJ0005, 1988).

6. Francis James Child, ed., *The English and Scottish Popular Ballads*, Vol. 5 (Boston: Houghton Mifflin, 1898), 88-95.

7. For example, the Bobby Rush recording discussed above is not without scholarly interest, but searches of OCLC and RLIN in June 1992 failed to identify a single library with a copy of it.

8. British Museum. Dept. of Printed Books. Thomason Collection, *Catalogue of the Pamphlets, Books, Newspapers, and Manuscripts Relating to the Civil War, the Commonwealth, and Restoration, Collected by George Thomason, 1640-1661* (London: British Museum, 1908).

9. See Hyder E. Rollins, ed., *A Pepysian Garland: Black-letter Broadside Ballads of the Years 1595-1639, Chiefly from the Collection of Samuel Pepys* (Cambridge: Cambridge University Press, 1922), and Roger Thompson, ed., *Samuel Pepys' Penny Merriments* (New York: Columbia University Press, 1977).

10. Elinor Des Verney Sinnette, *Arthur Alfonso Schomburg, Black Bibliophile & Collector: A Biography* (Detroit: New York Public Library and Wayne State University Press, 1989).

11. Gordon B. Neavill, "Electronic Publishing, Libraries, and the Survival of Information," *Library Resources & Technical Services* 28 (January/March, 1984): 76-89.

    12. John Tebbel, *A History of Book Publishing in the United States*, Vol. 1: *The Creation of an Industry, 1630-1865* (New York: R. R. Bowker, 1972), xiii.
    13.   Gordon B. Neavill, "Preservation of Computer-based and Computer-generated Records," in *Conserving and Preserving Materials in Nonbook Formats*, ed. Kathryn Luther Henderson and William T Henderson (Urbana-Champaign: Graduate School of Library and Information Science, University of Illinois, 1991), 45-60.