
JAMES H. SWEETLAND

ABSTRACT

While computerization has affected many library operations, it has had relatively little impact on most humanities work other than to make library collections more accessible within libraries (e.g., the OPAC) and among libraries (interlibrary loan systems). However, recent advances in technology are much more likely to have an immediate impact on arts and humanities work and its relationship with the library. The development of microcomputers and locally controlled laserdisk databases has already tended to a decentralization of control over information. This decentralization may cause libraries some problems such as those related to an increase in expenses and in the difficulty of bibliographic control. Electronic mail, by changing the way people communicate, may improve interdisciplinary contact, or it may create two categories of information rich and information poor scholars, as will technology in general. On the other hand, as the other articles in this issue of Library Trends indicate, there are many areas where machine-readable text and interactive computer systems will greatly assist the humanities scholar and possibly even improve general access to the humanities for the average person. In the short run, the most likely prospect is for all current systems to continue with electronic formats merely adding to the richness of the human record. In the long run, electronic information technology will very likely have important effects, but many of the changes will be so gradual as to be unnoticeable to those experiencing them until they look back.
When considering the humanist as a user of modern information technology, one must remember that information technology's fundamental design is based on the needs of the physical sciences and engineering. Even the name "computer" suggests the strong numerical bias at the start of the information age, a fact clearly evident from histories of information technology aimed at textual computing. Ironically, the earliest public demonstration, in 1960, of textual searching on a computer involved the *Golden Book Encyclopedia*, a general reference book. However, nearly all later work for years was dedicated to the storage and retrieval of scientific and engineering literature (Bourne, 1980). More recently a number of inventors and designers have worked farther from the industry's roots, but this distance tends to extend only as far as business applications. The graphics capability of modern microcomputers, such as the Macintosh, began primarily as a means of easing people into computer use; more sophisticated graphics began with CAD/CAM (Computer Aided Design/Computer Aided Manufacturing)—computer aided engineering—rather than art as such.

One obvious implication of this situation for the humanities is the dearth of relevant databases. The 1992 edition of *Computer Readable Databases* lists 6,383 databases of which 248 (about 4 percent) are classified as "humanities." On the other hand, the 1988 edition of this same directory classified only 2 percent of the databases in this category, suggesting the situation may be changing although slowly ("State of Databases Today," 1992). To put it another way, the humanities, although a growing part of the database world, are still a very small part of that world.

The arts and humanities have a number of characteristics relevant to the use of computer databases, whether these are very large remote systems, local library catalogs, personal or departmental systems, or any others. First, there is relatively little money available for the humanities and particularly for individuals. In 1988, for example, total federal funding available in arts and humanities was 296.3 million dollars, versus over 19 billion dollars for nondefense research (U. S. Bureau of the Census, 1990, pp. 237, 585). In part, this lack of funding is related to the perceived lack of need. The stereotypical writer, for example, needs only a cold garret, some paper, and ink to produce a masterpiece, while most science requires at least some equipment, and much modern science requires very large and enormously expensive equipment (Price, 1963). However, as the need for any level of technology grows, so does the need for money. The relation of technology's costs to money available is very important to any discussion of the future, because information and communication technology imply a need for money. If the individual
lacks the money to obtain a highly sophisticated workstation (which sells at present in the realm of $5,000-$10,000 for a full system composed of processor, high resolution color monitor, and necessary peripherals), then that person obviously cannot use it. Yet if total sales for such equipment are small, then the unit cost will continue to be high.

A second critical factor to consider is the humanist's need for considerable amounts of older material, a need very different from most other science and social sciences disciplines and most "practical" applications such as business. The extreme example is probably in the study of religion and philosophy where scholars still routinely study writings thousands of years old, some of which are available only in the original or in very few copies (Heisey, 1988).

Librarians are very conscious of the extreme difficulty in keeping humanities collections to a manageable size—each generation leaves an unweeable residue; the collection continues to grow. Regardless of the medium in present vogue, the humanities scholar will continue to use earlier technology for the foreseeable future. In practical terms, this means continued reliance on manual literature searches regardless of preference. Few chemists, for example, need refer to the pre-online Chemical Abstracts; few historians can avoid (even if they chose to) to use pre-1963 Writings on American History. For example, studies of citations in humanities fields show considerable use of older material, both primary and secondary, as opposed to other fields (Budd, 1989, pp. 13-15; Garfield, 1980; McCain, 1986). It is interesting that the recent Research Libraries Group analysis of information needs in humanities (Gould, 1988) found that all the humanities fields listed several types of retrospective bibliographies as a pressing need and listed rather few really radical electronic systems.

A third aspect of scholarship in many humanities fields is a heavy reliance on text. The actual words in various documents are very important. An abstract, no matter how good, only rarely can replace the actual text of a document. Imagine, for example, the number of poems whose content could be summarized as "nature is beautiful," a description of little value. This reliance on words used by a document's creator extends to the style and often to errors and corrections in different drafts and editions, so that different versions of the same document can be critical to many kinds of research (Weintraub, 1980, pp. 25-27). A related situation applies to visual "documents." Certain kinds of work (notably art and art history) rely very heavily on the use of extremely good visuals. The definition of "good" does not imply only an accurate representation of graphics; it often includes such things as the specific tone and value of given colors, the presence or absence of brushstrokes, and the like.
In brief, then, humanities research tends to require a large amount of material and often needs access to highly accurate copies of several variations of the originals if not the originals themselves. Thus, for example, a mere transcription of a document's text, no matter how accurate, may not be sufficient for a scholar. He or she may need to see the specific typefaces used, the page layout, and, perhaps, additions to the original document such as marginal notes. She or he may also need to see all graphics included with the document (Broadus, 1987; Garfield, 1980; Weintraub, 1980).

In addition to accuracy of reproduction, the earlier discussed needs clearly imply a requirement for large amounts of space. While electronic media are becoming ever more efficient, they are still far from the ability to handle practically infinite amounts of storage space. For example, a typed page, in retrievable digitized form, requires about 16,000 bytes of storage, but a full color visual of that same page requires about 1 million such bytes (Crawford, 1988, p. 200). The paper version of the color picture and the simple text, on the other hand, each require about the same amount of space with perhaps a thicker page for the picture (e.g., due to use of coated paper).

A fourth aspect of humanities work with very important implications for electronic information is the vocabulary issue (Wiberley, 1983, 1988; Wiberley & Jones, 1989). First, there is a general lack of standardization in terminology. Words may deliberately be used in more than one sense either because of the equivocation inherent in many languages or because the writer wishes to make a point. In fact, writers often will coin new words or use old words with slight changes of meaning because no current word precisely expresses the required thought. Second, words often change meanings over time, sometimes ending up expressing the opposite meaning from the original. This phenomenon applies to all use of words, but, of course, becomes most important when documents of considerably different ages are used, a common phenomenon in the humanities. A philosopher or historian may need to use very old material and will probably use documents more than a couple of generations old in any event, thereby increasing the chance that some words will have changed meaning.

A third aspect of this situation, of course, is that apparently similar words will mean different things in different languages, and sometimes even in the same language. This phenomenon becomes a particular problem when the writer or researcher must work in more than one language or dialect (SantaVicca, 1986; Winograd, 1984).

All of these difficulties with languages are particularly important because many retrieval and storage systems rely heavily on verbal
access. Most computer indexing systems do much of their work automatically, normally making no distinction among variant meanings of a homograph. Even human input systems, however, can run into the language problem, especially when translating from one language to another. The physical sciences and even business have been able to develop nonverbal systems such as the Standard Industrial Classification or chemical formulas in order to avoid some of the terminology problems. Many of the sciences have even been able to agree upon controlled verbal subject systems such as Medical Subject Headings (MeSH) while many humanities and arts subject fields lack even the rudiments of such a system. The continuing difficulty surrounding attempts to classify fiction merely for user convenience, suggests that more sophisticated work may be a long way off (Baker & Shepherd, 1987; Pejtersen & Austin, 1983; Shepherd & Baker, 1987).

A relatively recent development which addresses some of these issues is the Arts and Humanities Citation Index (A&HCI). While this approach has been common in the physical sciences since the early 1960s, the A&HCI only became available in 1978 (Garfield, 1980, p. 44). Rather than relying on either verbal or symbolic subject access, citation indexes operate on what could be called document matching. Since a given document tends to cite existing documents discussing the same topic, the citation indexes allow the searcher to avoid attempts to define the “aboutness” of a document. Rather, having identified a relevant document, the user merely asks the system for other documents which have cited the known one. It is interesting that relatively few humanities scholars appear aware of this database in either its print or machine-readable form and that its development was so late compared to the sciences. The latter was apparently due to the general lack of citation standards in many arts and humanities fields. The former may be due to the combination of lack of knowledge on the users' part and the fact that the current database indexes only cite journal articles while humanists tend to rely just as heavily on books and other monographs (Garfield, 1980). Note, however, that while the A&HCI is produced by computer and is available in several electronic formats, the true innovation has little to do with any electronic medium at all.

Another factor relevant to this discussion is the general lack of time pressure in humanities work. With very few exceptions, humanists rarely need to get the latest breaking developments as soon as they happen. The difference between humanities and sciences appears in two recent controversies. On the one hand, there is the case of “cold fusion.” Announcement of this discovery was made without even the usual peer review because of its timeliness (Mallove,
On the other hand, there is the case of the Dead Sea Scrolls where complaints about delay in their release finally led to a "bootleg" version after nearly forty years (Shanks, 1991). Much of the discussion about electronic media has tended to involve time pressures—electronic mail and journals will reduce delay in communication; online databases are better than CD-ROM because they can be updated more easily and quickly. In fact, the humanities have a tradition of accepting slower communications (compare the annual bibliographies in many fields with no intermediate updates) which reduces the attractiveness of the timeliness of electronic communication.

Humanists' need to search considerable bodies of text and a number of different texts also relates in a sense to time pressure. Given the lack of clear consistent indexing, searches often require considerable modification; the results tend to be rather large—several hundred records is not an uncommon retrieval set. Both searching and printing (or downloading) thus can take considerable time. This need is a significant factor in the use of nearly all remote online systems with pricing based on connect time online.

Research style is also a relevant factor. In contrast to the prevalent pattern in the physical sciences and a common pattern in the social sciences, most humanities publications have only one or, at most, two authors (Garfield, 1980; Stone, 1984, p. 294). In other words, to a high degree, humanists work alone. There is some evidence that some very informal collaboration does take place (Crawford, 1986, pp. 570-73; Hood, 1985), but overall the role of the proverbial invisible college appears very weak. Humanities work tends to be individualistic rather than collaborative.

Surveys of humanities scholars show that they tend to be uncomfortable with technology, although a significant minority are now using at least word processing software (Hirschheim et al., 1990, pp. 27-70; Morton & Price, 1989, pp. 33-45). Whether this resistance is merely unreasoning fear, or whether it is a perfectly reasonable response to systems poorly designed for humanists' needs is unclear. However, for the immediate future, the odds are high that many humanists either will not use much technology or will require some assistance at least in the details of its use. Given these characteristics of the humanist-in-general, several relevant trends in the electronic publishing industries are of great interest.

First is the fact that in some ways the industry is "maturing." This can be seen in the consolidation of business firms such as the DIALOG/VuText/Knight-Ridder and BRS/ORBIT/Maxwell mergers. It may also be seen in the increasing cooperation among firms such as the UMI/Wilson agreement to provide full-text articles from journals indexed in Social Sciences Index using UMI software,

At the same time, the laser disk database industry is growing more or less exponentially. Just five years ago, for all practical purposes, there were no laser disks available yet, by the end of 1990, Cuadra felt it necessary to add a new publication, *Directory of Portable Databases* (1990), which had over 950 listings. As in the case of the original online systems, most of these databases are in the business and scientific fields. However, the last two years have seen an increase in the number and variety of humanities and arts databases on CD-ROM and videodisks (Sweetland, 1991). In particular, much recent work in laser disk technology has involved interactive visual systems. From an early constraint of only a few pictures with motion or a fairly small number without motion, a number of laser disk systems now contain a rather large number of images complete with motion, stop motion, windowing, and other similar features. Even at present there are disks with adequate visual reproduction and at least some motion; the future should see great improvements in this area.

Remote online databases and laser disk versions, primarily on CD-ROM, are already a standard feature of reference rooms. Libraries are finding some major advantages to laser disk technology at least as it applies to "reference books." For one thing, the fact that it has a definite cost (as opposed to most online systems) makes the CD appear much like any other serial to the budget—i.e., the more a given CD is used, the cheaper the per-use cost becomes. This is much better for planning than an online system where the library, in effect, rents the information and so gains relatively little from increased use. Compact disks are also attractive for this reason. Having paid for a CD database, the library has almost no concern that a user is online to that database for over an hour other than the queuing problem of other users who might want to use the same database or the same equipment.

It is currently true that lines may develop, causing some libraries to restrict use—e.g., to fifteen minute segments—or require formal scheduling of CD systems. Of course such scheduling does put time pressure on users, although this is much less than in the case of connect time charges. The queuing problem results from the fact that present laser disks have generally required one workstation per database. Current developments in LAN (local area network) technology allow several workstations to be connected to a CD reader which can handle several disks at once. The simple version of this setup allows each workstation to access any database on the reader;
the more complex version permits multiple users to access the same CD (as it appears to the user) at the same time. The recent announcement by University Microfilms of a "jukebox" player holding up to 240 CDs and able to be connected with up to four other jukeboxes means the theoretical capability of handling over 1,000 different compact disks (Flanders, 1991, p. 858) and will probably be matched by other technology along similar lines.

While these approaches offer a solution to the multiple user problem, they can potentially lead to copyright and contract problems. With a central database accessible to several users at the same time, how many "copies" are being used? This question can be answered by the use of a site license, where the library pays some multiple of the basic price but is then permitted to allow a given number of users to access it at once. Regardless of the exact price of such a license, this does increase the cost of any given database. With a relatively limited budget, a library will be able to purchase fewer different databases as the cost of each rises. Since humanities databases tend to get less use, there is a danger that they will be those canceled or even ignored by the library.

Fortunately, other pricing options are available. H. W. Wilson announced in late 1991 that it would allow unlimited use by a given institution of a copy of its CDs (H. W. Wilson Company, 1991). This permission extends to all branch libraries of an academic or public system and includes remote access (e.g., by telephone). Clearly the firm is taking a serious risk that a group of libraries will not attempt to circumvent the obvious intent of this policy by forming a new megalibrary. But this announcement does suggest that at least one commercial firm feels it can make a reasonable profit without gouging its customers. If this sort of policy is adopted by other major database firms, it could solve many of the problems noted earlier. The danger here is clear: If potential sales of a given database are too small, there is little incentive to produce it unless the price is very high. Yet, if the price is high, only the most popular databases will be sold on the basis of the greatest good for the greatest number. Again, if the restriction is monetary, the odds are that the database not purchased will be in the humanities; yet lack of a customer base will mean lack of decisions to produce such databases. A pricing option which seems to address many of these concerns has been announced for OCLC's new "FirstSearch" system ("Special Report," 1991). This end-user online system requires a flat "per search" fee ranging from 90¢ to 45¢ (one command line is defined as one "search"; all connect time and printing is included in the fee). These sorts of options are very encouraging.

There is a major danger with any of these local database solutions, again connected with cost. An attraction of locally controlled
databases is that they avoid the unpredictable connect time charges of remote databases. However, since funds are always limited, there is a real danger that libraries will purchase copies of only most heavily used local databases, with a corresponding decline in the funds available for use of remote systems. This could be especially damaging to the humanities which tend to lack a single “best” database and to be very interdisciplinary. It would be a real tragedy if the advantages of laser disk systems eventually meant libraries offered unlimited searching in a few very popular medical, business, and physical sciences sources but had to charge fees or send users to outside agencies to search in the arts and humanities. It would be an even greater tragedy if the resulting decline in business led the remote online vendors to drop such databases entirely.

Money concerns are not the only problems. Another very real restriction on widespread acceptance of databases is the lack of compatibility. While hardware problems have been reduced through the development of standards, there seems to be an infinite set of permutations of keyboard plus computer plus monitor plus CD drive plus printer. A brief discussion with almost any librarian with experience in CD databases leads to the inescapable conclusion that not all these combinations will work, and that there is surprisingly little advice available from any source on what will work.

Even with the likelihood that hardware problems will work themselves out, user interfaces vary considerably. Some systems allow Boolean combinations; others do not or allow only “AND”. Some systems allow word proximity operators as well, but there is considerable variety among those that do. The exact commands needed also vary, with some systems requiring filling in a menu all at once, others allowing only one step at a time, and the like. Users can become frustrated very rapidly, especially when trying to use several different products in succession at the same workstation. The analogous situation would be if a person needed to learn different methods of reading with the books of each publisher. Since humanists are particularly likely to be frustrated with poorly designed or hard to use systems, they will be unlikely to tolerate libraries which force such systems on them.

While laser disks, as such, are of growing importance for the humanities, perhaps the more important technological development is “hypertext.” First appearing as Bush’s (1945) MEMEX and proposed in much its present form by Theodor Nelson (1974, pp. 44-45), hypertext has only recently become a reality for the average user. The various versions of hypertext allow more or less random connections among all parts of a document, different documents, or different types of documents. Rather than relying on any form
of indexing, a given user may compare different pieces of text by linking these and then displaying them. For example, a person reading a passage in a poem may go directly to a passage in a biography of the poet. Then, inspired by something in that passage, the reader may go directly to a historical account of the events in the world and in the poet's own city, plus a map of the city, and perhaps photographs of the poet's house and friends. This is possible even if no one had thought of this particular set of connections (Bevilacqua, 1989; Huston, 1990; Marmion, 1990).

Hypertext use is becoming quite common and will probably become even more so in the near future. Often connected with multimedia, hypertext systems designed for the humanist are now commercially available. Chen's "Project Emperor I," for example, includes text, graphics, and photos connected with the excavation of the first emperor of China and became available on CD-ROM in 1992 (Chen, 1991; Chen, in press). A more complex system, Project Perseus, includes considerable material on ancient Greece and was also available in late 1991 (Harward, 1988; Mangan, 1991).

These systems and others like them soon to follow have a number of important implications for both the library and its humanist users. First, hypertext effectively eliminates indexes as presently known. Given a body of text—or text plus graphics plus sound plus motion—a hypertext "shell" is attached to the text by an individual programmer. While the type and number of links attached to this text vary analogously to the access points created by an indexer, there are no terms from the text associated with these links per se. Or, in a very real sense, a number of programmers could create totally different hypertext systems from the same documents. More to the immediate point, however, having attached the shell to the documents, each user creates the links as she or he needs them.

As hypermedia documents become common, the library must consider their impact on access tools such as indexes. Traditionally, libraries first tried to obtain all relevant documents and created in-house access tools to them. As the body of literature became larger, the profession developed bibliographies and, later, collective indexes, first in print and then in machine form. Thus, at present, a library user has article level access to all journals indexed in RILA or the MLA Bibliography even though the local library lacks some of the journals. A growing number of libraries have begun adding at least some of these traditional indexing systems directly to their catalogs, for example, the Colorado Alliance of Research Libraries (CARL) (Lenzini & Shaw, 1991) and several NOTIS libraries have experience with online versions of periodical indexes loaded into their OPACS (Bakowski et al., 1990; Tenopir & Neufang, 1991).
The norm for research requires the user to rely on catalogs and bibliographies for rather broad access to the set of books likely to be of value, obtain these books, and rely on browsing or the books’ indexes for more detailed access. Similarly, article or chapter level access is available through indexes and abstracts. Many electronic systems allow selection of words or phrases in these access tools, which permits a greater chance of getting what is being sought but still limits searches to the text of the citation. If the full text of the document is online, it is possible to search the entire text, in effect allowing an expanded form of browsing. Hypermedia expands full-text browsing by providing links among sections of text or even among different documents. Any given term, word, phrase, or image can suggest a connection which the user can follow to another point in the same or different text or to another text indefinitely. While this permits the user to avoid preset connections or indexing terms, hypermedia access is not available at all without the text—lacking an index or catalog, one has absolutely no access unless one has the full text available. While it is clear that reliance on current catalogs provides poor retrieval (Lancaster et al., 1991), such systems as CARL or OCLC do allow users to find out something about documents not owned locally and not in machine-readable form. Even with nearly unimaginable hypermedia access, scholars will still need some kind of cataloging unless all publishing is done on nearly free generally available systems. However, even were this to occur, humanities scholars will still need access to older print on paper material, which is not likely to be converted into electronic form for some time. It is thus encouraging that experiments like CARL and NOTIS, among many others, continue.

Another aspect of hypertext with potentially significant implications for libraries is “repurposing.” There are already products on the market which consist not of linear text at all but rather a corpus of images plus text combined with a rather basic hypertext frame. The user of such a package can create a number of hypertext documents. For example, assume one library has a copy of Solar System Tutor (Softdisk Publishing) on a laser disk. Using Apple’s HyperCard, a reader could create a “slide show” about Halley’s Comet with graphics showing its orbit, photos of its appearance from earth, and a motion schematic showing it moving around the sun. She or he could then add such hypertext links as the capability to ask about cultural figures associated with the phenomenon. A user asking for such a link at this library might obtain a picture of Mark Twain (who was born and died in two successive appearances of the comet), along with some biographical information about Twain. However, given the same basic material, a user at another library might make
a similar link not to Twain but to Bill Haley and the Comets and a third link to H. G. Wells's *In the Days of the Comet*. A person who goes to library A and uses its system could obtain links to American literature. Later, that same person visits library B, presses the same "button," and finds not Twain but popular musicians. The concerned user writes a colleague who then goes to library C only to find references to British science fiction writers.

Is there, in any traditional sense, a "book" about Halley's Comet or an almost infinite number of such books? What happens to the presently accepted pattern of scholarship when neither a footnote nor a personal reference can be fully relied upon to point all readers to the same reference? What is the obligation of the library profession to provide either all possible links among such information or at least "bibliographic control" over them?

Access problems are potentially much more complex than even the foregoing suggests. Given the nature of microcomputers and modern software, almost anyone can generate a new document at any time with or without hypertext. For example, a number of developers produce HyperCard documents and advertise them in computer magazines or newsletters. Project Gutenberg (Graham, 1991) is trying to encourage entry of existing text into a national system. Nelson's (1991) Xanadu combines several of these trends, suggesting a system of "transclusion" in which a new "document" would not copy text from others but, would integrate portions of others only when the initial document was read electronically. The problems of bibliographic control in all this are staggering. Thanks to bibliographic systems such as OCLC, RLIN, and WLN, it is possible for many libraries to indicate that they own the same item. This capability becomes irrelevant if each item, thanks to its access points, is unique. While some systems (e.g., RLIN) have the capability to allow each library's specific records to display, the staff time involved even in entering "notes" for hypertext links could become staggering, assuming that the creators of such links would even tell the librarians in the first place.

Hypertext has the potential for changing the way scholarship is done (Lanham, 1990). For example, some schools of literary criticism emphasize the text itself in one way or another rather than the context of that text. Imagine a situation, such as the Perseus Project, where the text can be connected not only to other texts of a given writer, but also to background texts (e.g., primary and secondary historical works), pictures of the physical setting, and schematics (such as floor plans and maps). Imagine the possibility of connecting these elements to both music scores and actual sounds of the music of the time, plus perhaps reenactments of important historical events. Then
imagine that any reader can, within some broad limits, make any desired connections among these elements. In essence, hypertext could be the modern realization of the MEMEX but with even more capabilities than Bush considered (Nyce & Kahn, 1989).

The very nature of authorship may well change as a result of hypertext. Given the ease with which a “reader” can make connections among texts and graphics, the act of reading and the act of authorship may merge. The creator of the hypertext document chooses the pictures, plans, music, sounds, and text to be made available, setting up the original hypertext connections. However, the “reader” then may make new connections, even adding extra text analogous to marginal notes in a traditional linear document, an activity close to the author’s role. In a traditional written or printed document, the order of the text, arrangement of chapters, placement of notes and pictures, etc. are set once and for all when the document is printed. This is not true with hypertext.

In most of the experiments and proposals to date, the hypermedia package consists of a series of documents, essentially small “libraries.” Given the ease with which a reader could approach these documents, few users would even consider use of other documents in the library, certainly not those available only in printed linear text. To date, many such “libraries” have been put together by nonlibrarians just as documentary collections, textbooks, and the like have always been (although Project Emperor I, of course, is the product of a person with library training). Many library users, especially undergraduates, now prefer encyclopedias or handy collections of “major criticism” of “major poetry” while avoiding the rest of the collection even with the current state of affairs. Given this tendency to use the most accessible sources (Zipf, 1949), it is possible librarians will find a gradual drop in use of all printed documents followed by a decline in use of the more linear versions of the machine-readable texts. Given the enormous amount of documentary material available in world libraries and museums, such a trend would be most unfortunate to say the least.

Related to this is the question of collection development. There has been surprisingly little concern expressed about the implications of the purchase of such document collections. In effect, both the library and the individual abrogate their right to select material to the producer of the collection, much as they have done to the creators of current indexes. The tendency of many small libraries to refuse to obtain periodicals not indexed in Reader’s Guide to Periodical Literature is well known. This author’s own observations and discussions with a number of users of UMI’s Proquest business system suggests that many users and librarians have not even considered
its implications. While UMI does have the entire ABI/INFORM index in this system, it does not in fact have the full text of all journals in that index, only about 90 percent. From this author’s very unscientific survey, no one has really questioned the “selection criteria” leading to this result. Unless the profession is careful, even those users who do wish to go beyond “least effort” may be out of luck. Fortunately, there are the beginnings of discussions on this topic among the library profession (Hayes, 1990).

Another development of great importance even in the near future is the increasing number of large, interconnected networks with their related electronic mail and bulletin board systems. A surprisingly large number of libraries and colleges are already connected with each other, primarily through Bitnet and Internet. While the current systems leave much to be desired, they have many advantages besides rapidity of communication. For example, there is a growing number of library catalogs available (St. George et al., 1991; NSF Network Service Center, 1989). For another example, several humanities “discussion groups” already exist, although most are related to computer concerns. A major advantage from a financial viewpoint is that most use of this set of systems is free to the user, at least at this time. Thus one can “attend” a “discussion” in another state or nation without either the travel costs or even the telecommunication costs.

The “Uncover” experiment of the Colorado Alliance of Research Libraries (“On Target,...,” 1991) suggests a possible example of the near future for such systems. “Uncover” includes most of the contents pages of journals received in member libraries, readily searchable online. In “Uncover 2,” while online, the user may scan the contents and order copies of specific articles directly from CARL. Thus a library or an individual user anywhere in the United States can have not only bibliographic access at the article level but actual copies of the documents for a relatively nominal fee. In effect, Uncover provides an expanded version of browsing using traditional access systems. In a similar vein, the Research Libraries Group recently announced its “Document Transmission Workstation” which includes scanning hardware and software to permit transmission of any document between libraries.

The proposed NREN (National Research and Education Network) could help facilitate such activities. NREN could become the equivalent of the telephone system with everyone connected anywhere in the world with relatively simple use protocols. Fortunately, librarians and some humanists are involved in the development of this system which was originally considered only as a high-tech engineering science network. As presently proposed,
NREN would allow rapid access to a very large number of other networks and systems while requiring only one standard access language and standard protocols (Parkhurst, 1990). If it develops like Bitnet or Internet, telecommunication costs would be either free to the user or very inexpensive.

These sorts of national or international systems have the potential to eliminate the library as such. Humanities scholars tend to rely on personal collections, prefer browsing, and yet use many libraries. As of early 1992, it is now possible for the scholar to access thousands of library catalogs and order documents of potential interest at once limited only by budget. In theory, this could eliminate much of the role of the library and of the librarian. But, once again, the simple realities of money must be confronted. Even at a common per article fee of $10 per item and no other communication costs, the search for a few dozen items can easily equal the average scholar's monthly pay. For comparison, one could visit the Library of Congress and all other libraries in Washington for a week for well under $2,000 for all expenses. The same $2,000 would buy only 200 articles or chapters—fewer if the cost were greater than $10 each.

While laser disks and hypertext apply primarily to the local environment, electronic mail is already valuable in the larger world. Most common in business and the sciences, electronic mail is becoming popular with libraries. OCLC and RLIN both have rather sophisticated systems for use in interlibrary loan, and these were adopted rather early in the development of electronic mail, so in some ways we can say that libraries are a bit ahead of many other areas.

Electronic mail has several attractions for both libraries and scholars. It avoids telephone tag: a sender can compose a whole message, send it to the receiver, and be sure that the latter will receive the entire message as sent; the seemingly endless attempts of two parties to talk to each other can be avoided. The most important value of e-mail systems for the scholar, however, is their ability to foster communication among groups within the "invisible college." This can take place in two ways. First, a person can send a message to a number of people coded in such a way that each person's response to it is also automatically sent to all those who received the original message. Second, it is possible to set up a "bulletin board" so that anyone with access to the system can see the original message and all responses, making a response either to the original or to any of the others at any time. In effect, this resembles a conversation or a professional meeting. In fact, current systems are now being used to create exactly such conferences with some degree of success (Nunamaker, 1989).
There are several possible directions for this trend. One obvious possibility is the improvement of scholarly communication in the early stages or the expansion of the invisible college. As travel costs increase and as more people are involved in any given field it becomes more difficult for all interested parties to meet at annual conferences. If instead of, or in addition to, delivering a paper at a professional meeting the presenter also places it in a bulletin board, many who did not attend the meeting would be able to not only obtain the paper, but also engage in a dialogue with the original writer and with others who read the paper. It is possible that electronic communication will increase this sort of cooperative venture, thus facilitating the growth of invisible colleges in the humanities.

Electronic mail may also change the prepublication process. With more input from readers, the author could, in theory, have more information both on content and style before making a formal submission for publication. The entire refereeing process could also be speeded up considerably. The draft article or book, mounted on a restricted e-mail system, could be sent to reviewers for comments, which would be immediately available to the author and editor.

This process could extend to the actual publication of material electronically. To date, discussion of this option has tended to be in the context of speed of communication, particularly in the physical sciences and medicine. However, electronic publication could become quite important for the humanities by providing an "intermediate" form of publication between the article and the book. Given publication and distribution costs, scholars have found it progressively more difficult to publish short monographs (Thatcher, 1990). With its relative lack of expense, electronic publishing may replace the paper publication in this particular format regardless of its other effects.

Such systems could also eliminate the distinction between the publisher and the library (Rice, 1990). With a relatively small investment, libraries could provide access to referees and editors. Particularly in the humanities, libraries already hold much of the raw material of scholarship and produce bibliographic access to it. Adding the actual production of new material would only be a logical extension of these services. Presumably, if libraries became a significant force in publishing, prices and profits would fall and the emphasis would be on publication as access rather than as a source of profit. An early experiment in adding the library as publisher is taking place at Carnegie-Mellon University ("Publishing, The Next Generation," 1989, pp. 11-16).

This new medium has a number of problems, beginning with the requirement of access to the electronic mail system in question.
As with online and laser disk databases, each system requires different access protocols, and many e-mail systems require some financial outlay as well. In time, NREN may reduce these problems, but one will remain: Until a user logs on to a given system, she or he has no idea that there are any messages. This is roughly analogous to the situation if one were required to have a collection of mutually incompatible telephones each connected to a different group of people, and none of the phones rang, thereby requiring the user to pick up each one every so often to see if there were any messages.

A second danger relates both to costs and to the "information explosion." In many ways, the function of the library collection is to permit the user to avoid individual purchase of material through the collective purchase of potentially useful material. Assuming "publication" by e-mail (or any other remote electronic format) of any significant body of material, the library will be expected to obtain all material of relevance and download it for local access and for preservation much as it does with printed material. Since at least some of this electronically published material will be work in progress, the library may have to spend even more money than now for even more material, some of which will be of lesser quality (being in draft form). Again, since the sciences seem especially prone to creating new publications, the humanities could lose out in the competition for funds to download and store the material.

While a number of thinkers see the library's role as preserver of knowledge disappearing, the predictions sound much like what could have been heard in the late 1500s. After all, in theory, printing eliminates the need for libraries since anyone can easily obtain a copy of any new book assuming one has the money and has heard about the existence of that book. Strangely, printing does not seem to have had that effect. Similarly, mass availability of television has not led so much to the elimination of libraries as to the creation of new for-profit "libraries" called videotape rental stores, along with pressure on traditional libraries to add video formats to their collections.

As with lasers and hypertext, libraries will have difficulties with bibliographic control. Unless one posits that the library obtains everything from every electronic system, it will have to be selective in some way. The potential difficulty of finding out what exists could well be insurmountable. Not only will there be a question as to how to identify likely sources, but the question, What is the text? arises again. Since it is possible a given "document" will have comments attached over time, a given library's copy downloaded on Monday of one week, will be incomplete compared to another copy as it appears in the system by Friday. On the other hand, if the library refuses
to become involved in such material (as many libraries now refuse to become directly involved in the exchange of reprints, preprints, and conference papers), it may do a real disservice to scholars.

A case in point is an article that appeared in Academic Computing (Hallo, 1990) claiming stylistic differences (for the worse) in student papers written on Macintosh compared to IBM computers. The discussion was also mounted on a Bitnet bulletin board. As one might expect, both the printed article and the electronic version generated comments. Assuming one was aware of the Bitnet version and had the equipment and skill to download, one could obtain a very interesting discussion. While this material can be cited, there is no guarantee it will be available by the time the present article appears, so that the traditional value of citations for further research is damaged. Assuming the material was still available, by the time a reader of this journal looked for it, the total text of the original paper plus the comments will be different from those seen by this writer. As long as e-mail merely replaces telephone conversations, the lack of a permanent record is nothing new. However, if it is a form of "mail" or of "publishing," the value of the bibliographic citation and the role of the library as group memory have been affected as has the meaning of "publication."

Increased use of information technology may change the relationships of the librarian and the humanist. One very important aspect of that relationship is the often minimal role of the librarian. Generally, in contrast to the physical sciences, the librarian is rarely permitted any role as gatekeeper or information analyst. Many physical science operations have had such positions, and social scientists have at least discussed the issue (Line, 1971, pp. 425-26). Possibly the librarian might have a similar role in the humanities, although it is unlikely this would be as a "filter" as in the sciences. Rather, the librarian might become the guide and counselor through the maze of systems, command languages, and access points, not to mention the instructor in bibliographic control of not only the documents to be obtained, but also of those documents held by the individual researcher and those then created by the researcher.

The increasing introduction of automation into research and publication may well change the lack of collaboration considerably. Assuming librarians have some facility with technology and with explaining it in terms understandable to humanists, they may expand their role as "book-keepers" to include some form of bibliographic instruction. It is also possible that the ability to form invisible colleges, as well as this experience with librarians (not to mention programmers and system analysts), may habituate more humanities scholars to some type of collaboration. Of course this may not be
all to the good—the cliche about the camel being a horse designed by committee has some basic truth.

It is appropriate to close with a critical issue which has received surprisingly little attention—preservation. By now most scholars are aware of the rapid deterioration of much of our written heritage. In some ways computers may help deal with the crisis. Modern scanners and digitizers may help not only to preserve material which would otherwise be lost, but to preserve it in machine-readable form, thereby improving access while preserving the text. The quality of scanned images seems to be improving rather rapidly as is the speed of scanners. Many of the more important texts will probably soon be available in easily reproducible machine-readable form, some as part of a preservation project and many as part of projects requiring machine-readable records. One example of the latter is the *Thesaurus Linguae Graecae*. Now available in a CD form, this project began about twenty years ago primarily as an experiment in storing and retrieving text on computers. In the process of development, the compilers have expanded the total corpus of known classical Greek texts considerably (Watkins, 1991).

There is a negative side to this process, however. For one thing, it is not clear how permanent are many computer formats. Even the laser disk, contrary to early opinion, is susceptible to heat and physical damage. Present estimates give a laser disk a life span of only ten to thirty years (Crawford, 1988, p. 130; King, 1991).

The dangers to purely electronic media from the second law of thermodynamics and from magnetic fields are too well known and yet rarely discussed in the glow of all the new technology (Atkinson, 1990). It is possible librarians will expend considerable time and money to “preserve” in a format which will lead to its own crisis much sooner than acidic paper ever did. Present preservation techniques for electronic media recommend copying of all data on a regular basis—the copying restores the “freshness” of the electrons; use of a new piece of the medium avoids physical deterioration (Cribbs, 1987).

The odds are great that only the “most important” texts and images will be preserved and then only in the “most complete and accurate” versions. Since many types of humanistic research are based on variant texts—rough drafts, corrected editions and the like—librarians and archivists have a special obligation to be sure such material is also preserved. It is also likely that few repositories will be prepared with the staff and time to copy significant portions of their collection every year unless the library and archival communities make this issue an important part of their job.

Another potential problem relates to the preservation of the computerized documents themselves. Already there is more than one
generation of computer typesetting records, not to mention more recent word processing and data tapes. As authors and artists begin using the computer in the actual creative process, drafts formerly kept on paper are now only on computer diskettes. Preserving the diskettes, assuming the library ever gets them, is one technical problem, but how can the library be sure the author or artist saves the intermediate material? Few creators of documents consciously keep a copy of every version of every change. With the use of paper, retention of older drafts required only procrastination and a minimum of forethought. With electronic formats, the author must consciously make a new file for every new version, rather then merely saving the file under the same name, which is the customary approach at this time.

The development of machine-assisted information is still in its infancy. The long-term impact on humanists and libraries will be of the same order as the change to the codex from earlier formats or the change to printed documents. The main difference between the current changes and older ones will be the speed of the change. In the short run, it is most likely that many changes will be transparent and thus accepted by humanists with little trouble. For example, thirty years ago some seers predicted people would soon listen to music on computers. Today more and more people are doing just that—the CD player is, in effect, a dedicated computer. Similarly, word processors have become very common to the point where many journals accept or require submissions on floppy disk. On the other hand, most journals are still being produced on paper and many authors still use typewriters (some are rumored to prefer pens or pencils for that matter). In any event, it is well to remember the librarian’s “Fifth Law”: “The Library is a Growing Organism” (Ranganathan, 1964).

REFERENCES


Garfield, E. (1980). Is information retrieval in the arts and humanities inherently different from that in science? The effect that ISI's citation index for the arts and humanities is expected to have on future scholarship. *Library Quarterly, 50*(1), 40-57.


St. George, A.; Larsen, R.; & McKelvey, K. (1991-). *Internet-accessible library catalogs & databases* [Computer Database]. Albuquerque, NM: University of New Mexico. (e-mail: <stgeorge@unmb.bitnet> OR <stgeorge@bootes.unm.edu>)


