Between the Visionaries and the Luddites: Collection Development and Electronic Resources in the Humanities

Edward Shreeves

Abstract
The incorporation of electronic resources in the humanities into the traditional practice of collection development presents challenges that have pragmatic, technical, fiscal, and cultural dimensions. Many of the selection criteria valid for print resources have analogues in the electronic realm, while others are unique to the new medium. Among the most significant challenges will be understanding and responding effectively to the way computer-aided research in the humanities changes scholarship and scholarly communication. The substantial differences among humanist scholars in their readiness to participate in the evolving new world will require considerable attention from collection development librarians.

Introduction
The process of building a collection takes place within a cultural and social context from which it derives its values and assumptions. For selectors in the humanities, a part of that context has been formed by a system of scholarly communication and a literature that has developed over decades, if not centuries. As noted by many, that system is now suffering from severe strain on a number of fronts and is undergoing radical transformation. The emergence of machine-readable texts, of computer-based networks, and of all the attendant technological apparatus, has provided the means to alter radically scholarly communication and scholarly method in the humanities. Many librarians are eager to move toward this new future but are

Edward Shreeves, Collection Management and Development, University of Iowa Libraries, University of Iowa, Iowa City, IA 52242
LIBRARY TRENDS, Vol. 40, No. 4, Spring 1992, pp. 579-95
© 1992 The Board of Trustees, University of Illinois
unsure how to proceed. Both their eagerness and uncertainty have valid roots. A computer-based system of scholarly communication offers the hope of an escape from the insoluble morass of economic, space, and access problems of the last few years. There are even more reasons for uncertainty. The technology itself is in a state of constant flux with little probability for stability. The glimpses of the future, which now and then emerge from the mists, demand that libraries reconceive the ways they fulfill their missions or even redefine the mission itself. The price for this emerging system—both infrastructure and information—is likely to be very high at a time when many institutions are suffering their worst fiscal problems in years. Finally, changes in the social and cultural context in which scholarship occurs are taking place much more slowly than changes in technology. This article will consider some of the implications of these factors for collection development.

**THE ISSUES**

As machine-readable texts (of a kind that might interest scholars in the humanities) began to grow in numbers in the 1970s, there was considerable skepticism—to the extent there was any concern at all—about the library’s role in collecting and making these texts available. Some librarians, recognizing the research potential of these resources, argued that selection of research materials should not be limited by format (indeed, this argument was traditionally used for other nonbook resources), and that a computer file was simply another information package which libraries should collect. Like video recordings and microform, it had special features that differentiated it from print resources, but it was still an information source that supported teaching and research. In many ways this position, however enlightened and progressive, leaves unanswered a number of important questions and understandably did not foresee the world of networked resources that is haphazardly, but luxuriantly, growing today.

A concomitant argument deals with the issue of funding. In essence, this argument held that, while electronic resources were unquestionably useful, libraries should demand additional support to pay for them and not redirect dollars from already undersupported print resource funds. If librarians made a strong enough case to funding sources the argument ran, and local demand for these absent materials grew, the library would succeed in getting “new” monies to pay for these new and expensive formats. Although experience suggests that this tactic has rarely borne fruit, it is still heard in many quarters. It is tempting to assert that this argument, in fact, is not only fruitless, but perhaps dangerous, because it proffers an
excuse for inaction. As long as funding sources fail to provide the extra cash libraries need, there is no reason to look for ways to fund this new activity from existing resources. Taking the cost of electronic information from current resources is not a pleasant prospect, but it may be the only strategy available for many. A counter argument claims that such reallocation of existing resources gives funding sources a pretext for ignoring these new needs. This prospect leaves everyone at an impasse, and the potential beneficiaries—the humanists who need information resources—are the losers. It also puts the library at risk of abdicating its role as the organizer and provider of information for its clientele.

THE SCHOLARLY RECORD

In one way it makes sense to approach electronic texts in the humanities like their print counterparts. Books and journals are acquired to support teaching and research. To the extent that electronic texts justify the expenditure of resources, analogous selection criteria are valid. But many of the basic principles and practices of collection development assume the acquisition of an object—paper and ink or media carrying audio or visual information—which typically becomes a permanent part of the collection. A major function of the collection development librarian is to serve as a gatekeeper, identifying that portion of the published universe which a given library chooses to acquire. Selectors routinely perform this responsibility under a number of constraints. These include the availability of funds, the programmatic emphasis of the institution supported, the universe of publications and its accessibility, the number and skills of processing staff, and space availability. An effective bibliographer or selector should be familiar with the subject matter, including trends in research and publishing, knowledgeable about the strengths and interests of faculty and students, well informed about the book trade, and able to manage a budget. This bibliographer is judged, over the long term at least, by the collection he or she built—the aggregate result of specific, title-by-title, decisions made about which books (journals, microfilms, videos, etc.) to bring into the collection, and which to leave out. While many electronic texts can be purchased and acquired like their print counterparts, others, available through networks or through licensing arrangements, do not become a part of the library’s collection of information resources.

The theoretical and practical models developed for the processes of collection management have all shared this fundamental assumption—that the selector was exploiting limited resources to acquire that subset of the published universe most useful in the local
setting. Cooperation with other libraries collecting similar materials, with the avowed intent of avoiding unnecessary duplication and maximizing budgets, has received increasing attention, but the focus has always been the local collection. Present in the background has been the assumption that the cumulative resources of libraries and other repositories of textual information define what becomes the scholarly record. While the definition of what constitutes that record has been expanding in recent years to the point that virtually no source of information is outside the pale, libraries and archives, in the act of selection, effectively limit what enters the record. As Ross Atkinson (1990) points out, "[t]he definition of the record...has always been one of the library's primary social and epistemological functions" (p. 356).

Preservation is another concern. Various dangers to the physical integrity of collections pose risks to the intellectual integrity of the record and are a major reason preservation has received so much attention from collection management librarians. Collection development librarians have tacitly assumed that the information represented in the various media typically acquired by libraries would remain unchangeable and permanent. Atkinson (1990) explores the implications of a system of scholarly communication in which most information, textual and otherwise, being distributed electronically, is no longer immutable in the way print and other media are. He urges that libraries must continue to play a role in the definition and maintenance of the scholarly record. The method he suggests—to move “a carefully selected assembly of graphic utterances from the environment into a library database”—may not be the most efficient or desirable means to achieve the desired goal, but it does attempt to respond to the problem of record definition and preservation (p. 356).

The proliferation of discussion groups or lists on Internet and its affiliates raises some interesting questions for the collection development librarian. In the process of record definition and gatekeeping discussed earlier, there was little attempt to collect, except very selectively, the communications among scholars that took place prior to publication of finished products in peer-reviewed journals and books. The existence of these discussion groups, and the fact that the interchanges appear as text and are sometimes archived and searchable, has led some to wonder what role libraries should play in mediating access to them and preserving their contents. It is the invisible college made visible. Douglas Greenberg, in a paper delivered at the Symposium on Scholarly Communication held at the University of Iowa in November 1991, argues that “this is high-tech cocktail party conversation at a very high level and across very
great distances. Not all communication among scholars is scholarly communication." Greenberg's point is that only peer-reviewed scholarship, of which the networks presently offer little, represents real scholarly communication, and by implication, that electronic conversations among scholars, however interesting, do not form a significant part of the record and should not be a primary concern of librarians.

**Electronic Texts or Electronic Publishing?**

In discussing electronic texts in the humanities, it is necessary to draw some important distinctions, although these distinctions are not always equally relevant. Much of the interest in electronic texts among humanists so far has focused on those texts which make up the various canons of different fields. In some cases, like the *Thesaurus Linguae Graecae (TLG)*, this canon includes all known surviving texts in a particular field or subfield. In others, the body of texts is selective, though extensive, and the aim may eventually be to achieve some level of comprehensiveness. The *ARTFL* database of French literature and the Cetedoc patristic texts available from Brepols exemplify this kind of collection. In addition, there are many other projects of varying size and complexity and thousands of individual texts around the world which have been created to serve the immediate needs of a specific researcher or research project. In most cases the creator of these texts did not initially plan for electronic distribution of the actual text by way of networks or other means. The aim was typically the study of the particular text in question, the desire to produce a concordance, or some similar purpose. Scholars who note the unfulfilled promise of computer-aided hermeneutics (see below) are thinking of this kind of text.

By contrast, much recent attention on electronic texts, especially among librarians, has focused on the potential of a network-based system of scholarly communication as a replacement for the expensive and cumbersome world of paper and print publishing. In this brave new world, for which the article by Rogers and Hurt (1990) provides one model, the National Research and Educational Network (NREN) or its equivalent becomes a high-capacity pathway for distribution of scholarly information of all kinds under some combination of academic, governmental, and corporate sponsorship. Many view with eagerness the opportunity offered by such a system to wrest control of the scholarly publication from commercial sources. At the same time, it seems to hold out the only chance for libraries to escape from the endless rounds of cost increases which always overmatch budgets.
Under the scenario envisioned earlier, the system would allow for the dissemination of any "text"—e.g., a full-blown scholarly edition of the Greek poet Theocritus, an article-length critical study of his *Idylls*, a monograph about ancient pastoral poetry, or a hypertext commentary on particular poems. This kind of system would of course provide the same facilities for scholarly information in the sciences and social sciences as well. In fact, the economics of publishing in scientific and technical fields make such a system all the more attractive for that literature. The implications of such a future for collection development are profound. It could fundamentally alter the basic assumptions under which bibliographers and selectors have worked since libraries began to build collections.

To date, however, this system is still inchoate, and most electronic information in the humanities has been of a much more specific kind. As befits a collection of fields in which the text provides the essential object of study, most computer-based resources have been either bibliographic or what might be called canonical. A bibliographer can apply the selection criteria typically used in the selection of printed texts to these electronic counterparts. Many new issues arise, but some of the fundamental questions are still valid.

A crucial question, which has validity in both the electronic and print realms, has to do with the quality and authenticity of the text being judged. For printed resources, especially when dealing with often-studied texts, there are a number of ways to discriminate among texts of varying merit. Through reviews, consultation with faculty and subject expertise, the selector analyzes the reputation of the editor; the value of added material, such as an introduction, notes, and critical apparatus; the prestige of the publisher or series in which the text appears; and the care and accuracy with which the text is established. In addition, the work is considered in the context of a long tradition of published scholarship. Most electronic texts of the kind discussed here have, as source texts, printed editions which can be judged on these same grounds. It is not always clear what the source text is for a given electronic analogue, but most electronic texts are based upon an identifiable print counterpart. Their quality, then, may be judged, in part at least, on the merits of the original text. Depending on circumstances, it may be desirable to have a text in electronic format even if it is based on an inferior or flawed source, but the quality of source text matters.

Several electronic publishing projects illustrate different aspects of this issue. The *Thesaurus Linguae Graecae* has generally sought to base its texts on the best available scholarly edition of the work in question. The texts in Chadwyck-Healey's *English Poetry Full-Text*...
Database are derived from the latest edition published in the author's lifetime. In the case of the ancient Greek texts of the TLG, this approach was obviously not an option. But the English Poetry Full-Text Database, which avoids copyright issues by its chronological limitations and its choice of texts, may be providing less than ideal texts in those cases where a modern editor has published a version based on all the available evidence.

**Migne Versus Cetedoc**

The recent announcement of the imminent publication in electronic form of Migne's *Patrologia Latina* and of Cetedoc's texts in the *Corpus Christianorum* series provides a case study of issues that arise in selecting electronic texts in the humanities. These questions concern practitioners in the field as much as librarians, as some sharp debate on various electronic discussion groups (the *Humanist* and *Medieval Text—Philology, Codicology, and Technology* etc.) illustrates. The text of Migne's *Patrologia Latina* is being issued by Chadwyck-Healey, which is beginning to take an aggressive role in developing and marketing electronic resources of interest to humanists. The publication of Migne (as this collection of texts is widely known) was a major event in nineteenth-century patristic scholarship. Migne collected and printed what were purportedly the best available editions of the writings of the Latin fathers from the third through the thirteenth century. In light of the centrality of the church during this period of western European history, this collection became a standard source for medievalists in all disciplines. While many works contained in Migne's corpus have seen modern editions in the twentieth century, for a large number of authors and texts, the Migne edition is the only widely available source. Even when a new text has been edited and published, many scholars, for a variety of reasons, still cite his text.

At the same time, Brepols, a major Belgian publisher of patristics, has announced the publication of a CD-ROM entitled the *Cetedoc Library of Christian Latin Texts*. This collection will contain all the texts issued in the printed series *Corpus Christianorum, Series Latina*, and *Continuatio Medievalis*, more than 250 volumes, along with the works of Augustine, Jerome, and Gregory the Great. The printed series, intended eventually to replace and supplement Migne, presents up-to-date texts carefully edited to the highest standards.

The debate engendered by the announcement of these two publishing projects focused on a variety of issues including cost, the quality of the texts, searching capabilities, and tagging. It is perhaps notable that the greatest passion was reserved for issues of pricing since the Chadwyck-Healey product sells for around $45,000...
(the cost varying according to certain options), and many participants in the discussion found this to be unthinkably expensive. Much of the indignation seemed to be directed rather abstractly at the size of the sum—perhaps a year's salary for many humanists. It looks egregiously expensive in comparison with the very low costs of some electronic texts which have been produced by individual scholars or as a part of publicly funded projects. For example, the CD-ROM from the TLG containing all Greek literature is currently priced at $750 for institutions. In the case of Migne, the publisher has borne all the costs of keyboarding, proofreading, and tagging, as well as the costs of production and marketing, licensing a search engine, and maintaining the requisite corporate infrastructure. Public or privately donated funds have covered the costs for most of these activities for the developers of the TLG. The Cetedoc CD-ROM, on the other hand, is based on the machine-readable texts used in the production of the printed volumes which Brepols produces. In reacting to these issues, Timothy Reuter (1991a), pointing out the differences in the way the two products were developed, suggests that electronic texts will only become more affordable as the industry develops, but this will not happen "until a few of our institutions have shelled out megabucks for the earliest products." Bob Kraft (1991), another discussant, urges Chadwyck-Healey to take the product "back to the drawing board" and considered "calling for a boycott of the product by the libraries that are being threatened by this offer." Notable here—beyond an admirable concern for library budgets—is the tacit assumption that libraries will be the agencies acquiring these texts.

Many attacked the quality and accuracy of Migne's texts and noted that the Cetedoc editions were typically the most recent, and usually the best, texts available for the works in question. Defending the Migne product, others pointed out the value of having in electronic format a collection of texts which had been a standard source since its publication and still provided the only edition of many texts. Michael Sperberg-McQueen (1991) urged that "Chadwyck-Healey will do more for patristics, as well as all the other fields where PL [Patrologia Latina] is used, by reproducing Migne than by waiting another fifty years for better editions." Another discussant (Gaylord, 1991) worries that his librarian will buy the Chadwyck-Healey product because it is known as a standard work in medieval studies, and will not realize that scholars "who know what they are doing" are using more modern critical editions like those published in the Corpus Christianorum. This lack of confidence in librarians' understanding of the importance of textual authenticity is perhaps as worrisome
as the assumption that libraries will be the provider of such information.

Topics which received only some attention were tagging, searching interface, and networking capability. The Chadwyck-Healey products are among the first to be published using Standard Generalized Markup Language (SGML) compatible with the standards being promulgated by the Text Encoding Initiative (TEI). It is also being marketed as both a CD-ROM product and on magnetic tape and can be networked for multiple users in the same institution without additional cost. Several participants in the discussion spoke in glowing terms of the search engine for the Cetedoc product, but no one reported on the searching interface for Migne. Finally, one person noted that, whatever the quality of Migne's texts, what was more important was the accuracy with which Chadwyck-Healey had brought those texts into electronic form (Reuter, 1991b).

STANDARDS, SOFTWARE, AND CRITERIA

It is unusual to have a choice between two similar electronic resources like the Chadwyck-Healey and Cetedoc products. With the exception of Shakespeare and the Bible, there are relatively few duplicate texts to choose among. The Cetedoc-Migne debate is instructive because it mirrors, in part, questions that arise in any selection decision between two products that provide similar information, but also because it reflects concerns which could only arise in the online environment. Since others have discussed many of these issues and concerns, the following remarks will focus on some specific topics which seem to have serious implications for libraries attempting to collect electronic resources in the humanities.

In the humanities, particularly in those fields that focused on literature, history, religion, and similar areas where specific texts have great importance, bibliographers and selectors must know how to distinguish important texts from unimportant, authentic from unauthentic, scholarly from popular. Many electronic texts and text collections in the humanities, especially literary texts, are copies or versions of printed counterparts. Often the library already owns the published text which formed the basis of the electronic edition. The primary motive for acquiring the electronic version, as with the Migne and Cetedoc collections discussed earlier, is to provide users with the improved capacity offered by indexing and access software to access and analyze the text.

Several writers have pointed out that a major obstacle for the selector of electronic texts is the difficulty of defining the available universe. The usual selection tools (reviews in scholarly journals, national bibliographies, publishers' catalogs, etc.) do not cover such
resources effectively, nor is there a developed system of publication and distribution. Finding out about electronic texts requires attention to a number of specialized sources of information like the *Humanities Computing Yearbook, Computers and the Humanities, and Literary & Linguistic Computing*. The selector should also be familiar with the work of a number of projects and centers with a special interest in humanities computing, many of which publish newsletters or sponsor electronic forums. Examples from the English-speaking world are the Center for Electronic Texts in the Humanities, the Georgetown Center for Text and Technology, and the Oxford Text Archive. There are also a number of active centers in Europe and in other parts of the world. Gaunt (1990) and Lowry (1990) describe some of these, and other, resources, but no one would claim that their coverage of humanities computing is yet either systematic or exhaustive. Scholars and librarians share this problem. One of the most common inquiries on the Humanist Discussion Group takes the form, Does anyone know of an electronic version of some literary or historical text? Within libraries it is collection development librarians, with their subject and language expertise, who are in the best position to keep current with this chaotic situation and make informed judgments about resources.

Lowry (1990) distinguishes between published electronic texts, defined as those intended for further distribution—and unpublished texts—those not intended for further distribution (p. 16). This distinction, in what has until recently been mostly a scholarly cottage industry, seems largely unnecessary. Many of the sources which now actively distribute their texts or corpora began as projects with limited aims and little or no thought of distribution. In part, it is the microcomputer revolution and the growth of cheap mass storage capacity, in part the growing audience of users that have made distribution feasible. What it means to publish an electronic text, in view of the ease of duplication and dissemination and the informality typical of the process, is not at all clear.

As mentioned, many of the criteria governing selection of printed materials—the intrinsic importance of the text, the care with which it was established, its pertinence to local needs, its relationship to the existing collection, its cost—retain validity in the online environment (Johnson, 1990, pp. 7-9). There are also characteristics unique to electronic texts to complicate the task of the selector. These include criteria related to markup or tagging, to access software, to equipment platforms, to standardization generally, and to incorporation within standard library practice.

One cause of the seemingly chaotic situation with regard to electronic texts stems from the difficulty of finding out what is
available and how to get it. Another results from inadequacies in the existing systems for encoding characters and marking up text. Two recent efforts, Unicode and the Text Encoding Initiative, have sought to address these problems, and their success or failure could have significant implications for the spread of electronic resources in the humanities.

The inadequacy of ASCII encoding schemes to represent all the world’s languages is widely recognized. Many feel that, for humanities computing to reach its potential, it is essential to find a single solution to accommodate the multiplicity of character sets and scripts used around the world. The Unicode initiative has emerged as one attempt to meet this need. Unicode proposes a single encoding scheme for all currently spoken languages, including those in non-Roman scripts. Unicode proposes a 16-bit (compared with ASCII’s 8-bit) coding scheme, with a capacity of 65,536 codes. Its developers hope that its adoption will make it simpler to write multilingual software and exchange information worldwide. It is no accident that a library consortium, the Research Libraries Group, is one of its many corporate sponsors. Libraries, especially large research libraries, seek to represent in their online catalogs bibliographic information about items in nearly every known language. The failure to date to represent East Asian vernacular scripts (among others) in existing catalogs reflects how far from the ideal are these tools (Elman, 1991). Unicode may not turn out to be the successful solution to the multiscrypt problem, but it does represent an encouraging move toward standardization and interchangeability. It remains to be seen if centrifugal market forces will inhibit or prevent the acceptance of a standardized basis for multilingual computing, whether Unicode or its surrogate.

In a sense, the Text Encoding Initiative seeks to do for the text what Unicode attempts for the character. TEI aims to produce what some have called a metalanguage—i.e., a coding scheme which will enable texts to be created in a standardized form transportable from one hardware and software platform to another without loss of information about structure and textual features. In an electronic text, many of the usual signs which communicate the structure and organization of the text are peculiar to the computer system under which the document was created—or in many cases completely absent. Based on SGML, the TEI has produced draft Guidelines (Sperberg-McQueen, 1990) with the intent to support data interchange and application-independent local processing and to offer guidance in text creation or capture (p. 1). Funding received from the National Endowment for the Humanities and the Commission of European Communities indicates the potential importance of the TEI for
humanities computing. Libraries have based their collections over the centuries on the fundamental, almost unexamined, assumption of the interchangeability of printed texts. If the TEI succeeds in meeting its goals, it could form the basis for a more coherent and orderly growth in the development and use of electronic texts.

One obvious feature which differentiates an electronic text from a printed one is the necessity for equipment and software. The need for equipment is analogous to the situation for a videorecording or microfiche, but there are greater complexities involved. In addition, the situation is more unfamiliar because, until recently, most selectors were not accustomed to factoring equipment considerations into their decision-making process. Electronic texts in themselves have limited utility. To manipulate these texts requires software that searches, displays, and otherwise manipulates the text in ways that serve users' needs. Some text files are only accessible through tailor-made software, while others may be used with a variety of packages. But whatever the interface, the utility of a given text file might depend more on the availability of Micro-OCP, FolioViews, or Wordcruucher software than on the quality of the text itself. The quality and accuracy of this software must receive greater attention from scholars and librarians. It is essential that it really do what it claims to do, but there are fewer analogies to the world of printed resources than in the judgment of electronic texts themselves. This dependence on unfamiliar equipment and access software imposes new requirements on both the selector and those who catalog and provide service for such resources.

VISIONARIES AND LUDDITES

Greenberg, at the Iowa Symposium on Scholarly Communication, described two extremes in the scholarly community, between which, by implication, fall the vast majority of working scholars in the humanities. The visionaries, like many librarians who have eagerly embraced technological solutions, "believe that the emerging technologies signal a radical change in every aspect of scholarly communication" and are impatient to get on with the transformation. At the other extreme, the Luddites are "unalterably wedded to print" and have no use for online catalogs much less online texts of Shakespeare. The visionaries need no convincing that the future of scholarly communication in the humanities lies with computer-based technology. The Luddites are probably immune from persuasion for the most part, but when they are influential senior professors, their opinion carries considerable weight. At the same conference, in another context, a presenter suggested that the only
solution to the barriers thrown up by Luddites, especially in positions of authority, was to bide time until they retire or die off.

One of the great obstacles facing the academic librarian eager to provide access to electronic texts, especially in some humanistic fields, is the hostility of the Luddites and the indifference of the silent majority in the middle. While a small and growing band of enthusiastic visionaries does exist, the numbers, if not the intensity of their interest, are small. As Erwin Welsch (in Bailey & Rooks, 1991) recently pointed out, a comparison of the number of participants in the Humanist Discussion Group (over 1,200), a well-established, international moderated list server for humanists, to the number subscribing to PACS-L (over 3,000), a list server mainly for librarians, illuminates the degree of difference between levels of librarian and faculty interest in electronic communication (p. 30).

Most scholars are neither visionaries nor Luddites and it is this large and largely indifferent group that must be convinced of the efficacy of technological solutions to the problems of scholarly communication. When confronted with the likelihood of seeing fewer books and journals in order to pay for sometimes very expensive electronic resources, faculty in this group understandably balk. They urge that libraries should not be experimenting with new electronic gadgetry when they are having so much trouble responding to the existing demand for print. This line of reasoning is especially resonant when libraries are gutting subscription lists and acquiring fewer and fewer monographs.

Their resistance produces a quandary for the librarian. On the one hand, many have articulated a clear responsibility to lead the way in pointing to the benefits of broad access to electronic texts and helping the uninitiated find what they need and learn how to get it and use it. On the other hand, there is the risk of alienating the goodwill of strong library supporters by getting too far ahead of them, especially when they see human or material resources diverted to this end. Yet it would be irresponsible to ignore the potential which computer files have for the processes of scholarly communication and the analysis of text and image. Besides, there is the ever-present risk that other agencies may expropriate the role of libraries. If this is the case, librarians must continue to build alliances with those faculty who are users of electronic resources and work to convert the nonbelievers through education, demonstration, and experimentation.

One reason for the indifference of this silent majority is the absence of clear evidence of the benefits of a new technology based order. Faculty at most colleges and universities, at least in the United States, have eagerly embraced the microcomputer and some of its
associated software, chiefly word-processing systems, but few have ventured far beyond. Even though many scholars of language and literature study texts in ways that could benefit from computer-aided analysis, they have not exploited this potential for a variety of reasons. One has been the lack of easily available texts and of easily used and widely available software. Many are simply unaware of the kinds of questions that can be answered by electronic texts or are reluctant to spend the time required to learn how to pose such questions through the computer. Still, as Eric Dahlin (1991) points out, "humanities computing is one of the fastest growing fields of computing at present" (p. 4) and this rapid growth promises to speed the rate of acceptance of technological approaches to scholarly work in the future. But there remain many inhibiting factors.

Observers have often noted that the scholarly reward system—i.e., the tenure and promotion process—does not encourage the creation of humanistic databases, the development of software to manipulate them, or the publication of scholarly research based on electronic texts. At the same time, the market for such products in the humanities has so far been too weak to support the kind of commercial investment required to produce scholarly tools. At a "Computer Files Workshop" sponsored by the Research Libraries Group on January 11, 1991 in Chicago, Mark Olsen, assistant director of the Center for Information and Language Studies and the ARTFL project at the University of Chicago, maintained that there has not been a shift in perspective as a result of computer-based resources in the humanities as there has been in the social sciences with its computer-based use of quantitative information. He noted that the increase in the number of humanities databases does not seem to have produced a proportionate increase in the amount of research based on them. Faculty often report that, while their interest in creating and working with electronic information sources is keen, their mentors urge them to produce traditional scholarship for publication in mainstream journals if they want to be eligible for tenure. At present, most research based on electronic resources is published in marginal journals, which count less when tenure review time approaches. Certainly contributions to most online discussion groups and even to the few refereed electronic journals that do exist do not carry the prestige of an article appearing in a leading print journal.

Olsen and others have suggested that the problem is more than simply a matter of insufficient credit awarded to those working in the area of humanities computing. In an announcement of a session held at the 1991 Modern Language Association convention appears the summary of a paper to be given by Olsen which seems to develop
these themes. According to the summary (McCarty, 1991), Olsen argues that

the computer has not had significant influence on literary studies because old models remain dominant. He emphasizes that the primary benefits of the new tool come from asking new questions with it, but that first we must construct an appropriate model of computer-assisted research based on what the machine is particularly good for. He notes that there has been little interaction between critical theorizing and computer programming, to the detriment of both, and recommends concentrating on the specific theoretical and methodological issues.

Thomas N. Corns (1991) comes to a similar conclusion when he notes that "there is no substantial body of achievement in the field of computer-based literary criticism in English studies, and a discipline that has hardly begun can scarcely be a discipline in crisis" (p. 127). He supports this pessimistic conclusion by an analysis of papers published in the specialist journal Literary & Linguistic Computing (LLC) and several nonspecialist periodicals. Even in LLC (and its predecessors) the amount of computer-based literary analysis has been fairly small, and gotten smaller over time, while in a limited sample of four mainstream journals only one article (a computer-based analysis of prose style) appeared in the entire decade of the 1980s. Corns recalls that previous acknowledgments of the failure to produce significant results had found reason for optimism in future potential. In reaching his conclusion, Corns, like Olsen, blames the lack of results on deficiencies in the theoretical underpinnings of computer-aided analysis. Corns suggests that the concept of intertextuality can provide a useful theoretical base for computer-assisted work in which large databases of properly encoded historical texts help provide to students and scholars the intertext that would have been available to contemporary readers. It is certainly ironic that Corns's response to the problem of unrealized promise, however valid, is another appeal to the potential of the future. If librarians have a role in addressing this theoretical failure, it will be to act as facilitators providing scholars with the resources necessary to develop and test new theoretical models.

At the same time, there is evidence that technology can change the way scholars do their work. Theodore Brunner, director of the TLG, claims that this project provides "a prime example of how a humanities discipline has changed fundamentally for the better in consequence of the acceptance of technology" (Watkins, 1991, p. A24). Perhaps because the field is defined by a finite and relatively fixed set of texts, perhaps because it began before many other projects, perhaps because of the low cost of the CD-ROM version of the database, the TLG has made a significant impact on the kinds of questions scholars in classics can ask. But many of the problems
which the TLG was able to solve on the way to transforming research in classics are not as tractable for other areas of study.

CONCLUSION

The preceding paragraphs have sketched some of the issues which must concern the collection development librarian in determining the role of computer files in building collections as well as in making particular purchase decisions. While technical concerns and questions about standardization and interchangeability demand attention, the selector of electronic resources in the humanities must start with a knowledge of the subject matter, the methods and issues of the discipline, and the needs of the local program. This knowledge, equally important for selection in any format, is all the more necessary when selecting in an area marked by constant change and ambiguity in which personal knowledge and informal networks are often the best source of information. This is an argument for training the subject specialist in the requisite technical knowledge rather than seeking to educate the technologically adept in the appropriate subject disciplines.

The effects of electronic texts in the humanities are likely to be profound, and the strength of the impact will surely increase at a geometric rate in coming years. Many of the challenges facing librarians, and especially collection development librarians, will be technological and economic. But the greatest challenges will be social and cultural as selectors face the need to transform their own basic assumptions and to take a role in changing the ingrained views of faculty. It will be essential to make alliances with those visionaries already converted by the promise of computer-based models and work with them to demonstrate the validity of these models to an unconverted and uncaring majority.

REFERENCES


Reuter, T. (1991b). Further responses on Migne. *Humanist Discussion Group*, 5(324). (Electronic Discussion Group, e-mail: <humanist@brownvm.bitnet>)


Sperberg-McQueen, C. M. (1991). On Migne. *Humanist Discussion Group*, 5(0311). (Electronic Discussion Group, e-mail: <humanist@brownvm.bitnet>)
