Libraries and the Creation of Electronic Texts for the Humanities

Natalia Smith and Helen R. Tibbo

The creation of electronic texts (e-texts) in libraries presents a vision of the future for both libraries and the humanities. With the technological evolution on campuses today, many academic libraries are becoming producers of information as well as filling the traditional role of being archival repositories. A well-planned and articulated alliance among the university library, the computing center, and the faculty can result in the successful implementation of e-text projects and centers for individual campuses and for the networked scholarly world beyond, but this is far from a trivial undertaking. The primary goal of this article is to produce an understandable framework of the issues and challenges facing librarians as they consider implementing e-text projects.

Although it sounds like a cliché, we are living during an information revolution and are witnessing a true paradigm shift in how information is viewed and managed.[1] As the federal government, corporate executives, and others discuss the dizzying speed of today's information networks, James H. Billington, Librarian of Congress, talks about the key role of the nation's libraries in this process. According to Billington, "if we didn't already have libraries, we'd have [to] invent them" as libraries "are the keys to American success in fully exploiting the information superhighways of the future."[2] He argues that if the new digital media are truly to serve America, they must do more than offer entertainment ("infotainment") or provide information on demand, often too expensive for the majority of Americans. Unless we are to have a nation of "information 'haves' and 'have nots,'" libraries must "participate fully in the design, construction and maintenance of America's new information infrastructure."[3] Libraries are not the only essential, but chronically underfunded, institution concerned with the information superhighway; the humanities also struggle to fulfill their role in the development of this communication medium. A 1994 report from the Getty Art History Information Program, the American Council of Learned Societies (ACLS), and the Coalition for Networked Information (CNI) observed that:

> The absence of the humanities and arts in the development of a national information infrastructure ignores the value of the American

Natalia Smith is a Digitization Project Librarian in the Davis Library at the University of North Carolina at Chapel Hill; e-mail: nsmith@email.unc.edu. Helen R. Tibbo is an Associate Professor in the School of Information and Library Science at the University of North Carolina at Chapel Hill; e-mail: Tibbo@ils.unc.edu.

people's cultural heritage, and the network as a medium for creativity and learning, in the crucial formation of technology policy.[4]

This report goes on to point out that "reinstating the humanities and arts in the dialogue shaping this public policy is of utmost urgency," and calls for "the reintroduction of the humanities and arts in the formation of such policy." The prevailing attitude in Washington regarding the elimination or downsizing of agencies including the National Endowment for the Humanities (NEH) and the National Endowment for the Arts (NEA), however, casts doubt on what type of influence humanistic scholars and artists will have on the development of the Internet and the resources to which it will provide access. With today's new digitizing technologies, libraries, at least our large research libraries, may come to the aid of humanistic scholarship by assuming a publishing role as well as their traditional archival function. Although producing print materials has always been more costly than libraries could afford, the conversion of previously published and manuscript materials into electronic formats is now a possibility. Dividing the labor among libraries and sharing the texts across networks makes the idea of library-produced electronic versions of humanistic texts feasible, yet the amount of work involved should not be underestimated.

This is yet another example of how science and technology is so often well supported while humanities projects languish for lack of funds.

The creation of electronic texts (e-texts) in libraries is a very recent phenomenon, but already many libraries and similar institutions have started their own e-text projects. The Digital Library Initiative (DLI), funded through the National Science Foundation (NSF), the National Aeronautics and Space Administration (NASA), and the Advanced Research Projects Agency of the Defense Department (ARPA) from 1995 to 1999, will provide $24.4 million to six projects designing, developing, and testing elements of a digital library.[5] Thanks to those grants, research will be conducted on a significant scale in a distributed environment, with the emphasis on sources accessible via the Internet. Although many of the DLI-sponsored projects will provide the foundation for how to store and transmit digital information, their focus will remain primarily in scientific and technical areas. This is yet another example of how science and technology is so often well supported while humanities projects languish for lack of funds. Humanities materials pose significantly different digitization challenges that also will need to be studied and supported if large quantities of humanistic materials are to be available in a networked environment.

Fortunately, some work with digitizing nonscientific materials is going forward at the national level. In the fall of 1995, the Library of Congress announced its CANDLE (Creating a National Digital Library for Everyone) project. This program is designed to provide schoolchildren with electronic materials, help them integrate these sources in their studies, and develop critical thinking skills. Billington explains that rather than replacing books, "electronic collections are bringing people back to books."[6] Bell Atlantic is giving the Library of Congress $1.5 million for the CANDLE project, one part of LC's larger National Digital Library (NDL) Program for which it has raised $18 million to date, with a $60 million goal by the year 2000. Despite such projects, we are only at the beginning of a long and potentially painful period of transition rather than at some promised land. With the exception of recent scien-

tific data, most recorded human knowledge still does not exist in digital form but, rather, in myriad print and manuscript formats.

Humanists and Information Technologies

Information technologies can improve humanistic research in many ways. By taking "the drudgery out of lengthy philological, linguistic, or text analysis while providing a measure of accuracy normally unattainable by human effort alone," computers can greatly speed up these processes and make them more accurate.[7] Because computers can look at so much text so fast, text analysis programs make it possible to produce more reliable, more thorough results in less time.

For many years, humanists who wanted to use computers to analyze texts to look for patterns and other stylistic features had to create both their own text analysis software and the e-texts themselves. Much of this activity and struggle for scholarly acceptance has been documented in the journals *Computers and the Humanities*[8] and *Literary and Linguistic Computing*,[9] and the Humanist listserv.[10] These factors have meant that adopting new, technology-based research methodologies, such as e-text analysis, has been a very risky proposition, especially for nontenured faculty, who have little time to devote to programming activities or any work that colleagues will not readily accept.

Casual observation and survey evidence show considerable change in the nature of humanities research today. Many humanists now recognize the advantages that e-texts offer for both teaching and research. Machine-readable texts, easy-to-use text-processing packages, locally constructed databases, and national networks are changing humanists' attitudes toward technology, their communication patterns, the way they do research, and the way they use information. Scholars have discovered that computers can process both qualitative and quantitative research data. In the field of literature and linguistics, computers can easily recognize patterns, analyze text and style, and model concepts. As a result, the humanistic disciplines are becoming increasingly computer dependent, and e-texts are at the heart of much of this work.

A New Vision for Libraries

To maintain the support of some of their most frequent users, libraries must adapt and respond to the new electronic environment and its possibilities for humanists. Appropriately, more and more academic institutions are opening e-text centers that will serve their campuses and the networked scholarly world beyond. As of May 1996, the Center for Electronic Texts in the Humanities' (CETH) *Directory of Electronic Text Centers* indexed twenty-six institutions in the United States.[11]

Katharina Klemperer (assistant director for systems development, Harvard University Library) rightly compares e-text practices in most libraries today with what we could have observed in the early 1980s with OPACs (online public access catalogs): "just as the pioneering OPAC institutions were then starting to convert their card catalogs, the electronic text pioneers . . . are now starting to build their own collections, often 'converting' and marking up the texts on their own."[12] This endeavor is particularly important for works that electronic commercial publishers would be likely to overlook. For example, although excellent electronic versions of Shakespeare's corpus may be accessible, it is far less likely that

more and more academic institutions are opening e-text centers that will serve their campuses and the networked scholarly world beyond.

nineteenth-century testimonial materials from the American South will be widely available in a commercial product. To meet this need for both local and remote scholars, and to continue to maintain its excellence in southern Americana in general, the University of North Carolina at Chapel Hill is presently creating just such a collection of e-texts.[13]

Recently, other important text-encoding projects have started within, or have been directly associated with, libraries. This development seems to be a logical and essential course within general technological and information delivery transformations in our society. As David Seaman points out, "an electronic text initiative belongs in a library because it is a textual as much as a technical endeavour, although the two are sometimes difficult to separate. Certainly, the textual, bibliographic, and educational skills needed to evaluate, prepare, and present electronic texts to users unfamiliar with such services are all found in libraries."[14] Seaman successfully implemented this idea in the Electronic Text Center at the University of Virginia at Charlottesville.[15] Other interesting text projects reside at a number of university libraries, including the University of Michigan,[16] Georgetown University,[17] and Indiana University.[18]

In 1990, responding to rising voices in the humanities community, Princeton and Rutgers Universities wrote a joint proposal to the NEH to create CETH. The center was established in 1991.[19] From its beginning, "CETH was envisioned as a national node on an international network of projects and centers actively involved in the creation and dissemination of electronic texts to support humanities research and scholarship."[20]

One of the new center's major goals was to provide centralized control over a growing pool of existing and planned e-texts. As Susan Hockey, permanent director of CETH, stated, "the picture in the early 1990s [was] one of many humanities texts in many different places and in many different formats."[21] Librarians and scholars simply did not have any opportunity to keep track of existing collections all over the world. The situation was becoming troublesome for everybody, as "there ha[d] been few systematic efforts to make existing electronic texts available for other scholars to use,"[22] with a notable exception of the Oxford Text Archive (OTA).[23] CETH and other smaller text production centers and archives are responding to the need for e-texts in the humanistic community.[24]

Important Issues in the Production of Electronic Texts

Libraries are slowly augmenting their traditional role of serving only as archival repositories of information and gradually starting to act as producers of information. In the past, this has been prohibitively expensive for most printed materials. The technological evolution on campuses, in combination with a three-way partnership among the library, the computing center, and the faculty, offers a new potential for creating university publishing enterprises. However, this sort of enterprise will become a reality only if based on a well-planned and articulated alliance among these three forces.

As librarians consider creating their own e-texts in-house in addition to selecting commercially available products, they should first conduct relevant user studies, undertaken in the library setting, to see if there is sufficient academic need to warrant the development of "customized" e-text products. If there is, the staff should create a collection development and management plan focusing on e-texts. The plan should include not only policy for selection of texts, but also a clear framework for the production and maintenance of these materials and the evaluation of the project as a whole. Finally, the plan and the development of networkable texts depends upon a respect for standards and an understanding of the need for quality control.

The last, but not least of the factors mentioned above—standards and quality control—is perhaps most important as the number of e-texts continues to grow. Unfortunately, we see numerous e-text projects in both the commercial and academic communities that conform to no standards and exhibit little or no quality control.

Because of a lack of quality control, today we face a pool of e-texts that "vary greatly in their accuracy and the features which have been encoded."[25] According to the OTA, "some have been proofread to a high standard, while others may have come straight from an optical scanner. Some have been extensively tagged with special purpose analytic codes, and others simply designed to mimic the appearance of the printed source."[26] As John Price-Wilkin notes, "because of the cost of creating the texts, investing in the texts must be an investment in the future."[27] That means that e-texts should meet high standards of quality. An urgent need for developing text-encoding quality guidelines was discussed at the ACRL E-Text Centers Discussion Group at the ALA Midwinter Meeting in January 1996.

Today, the cost of digitizing materials is too high to do it just so we can say that we have a digital information project under way. The digitization process usually includes keying or scanning (or, in some cases, some combination of both), proofreading, imaging, coding, and creating a database. Judith A. Zidar, coordinator of the National Agricultural Text Digitizing Program at the National Agricultural Library, stresses that the most significant cost goes for "the editing, correction of errors, and spell-checking, which though they may sound easy to perform require, in fact, a great deal of time."[28] The coordinators of the American Memory Project of the Library of Congress (now a part of the Digital Library Project) had a happier experience with finding cheap labor. As a government agency, the Library of Congress contacted another government agency—Federal Prison Industries—and negotiated a contract beneficial for both sites. According to Ricky Erway, associate coordinator of the American Memory Project, the Library of Congress "shipped them [the prisoners—N.S.] the books, which they would photocopy on a book-edge scanner. They would perform the markup on photocopies, . . . perform the keying, and return the material to AM on WORM disks."[29] The results of that joint venture were approximately three dollars per page, without imaging. However, most libraries will be unable to employ prison labor, and current estimates for text conversion range to six dollars per page of text.[30] These figures show how carefully and thoughtfully libraries should approach the problems of creating e-texts.

The authors conducted the survey discussed in the remainder of this article in order to produce an understandable framework of the issues and challenges facing librarians as they consider implementing e-text projects. Almost all the issues associated with the production and maintenance of e-texts can be grouped into six categories:

1. text selection;
2. data conversion;
3. data presentation, including markup schemes and level of encoding issues;
4. access;
5. archiving;
6. user and staff education and training.

The authors' findings from a survey of some of the pioneers involved with library e-text projects present data in these six areas.

Survey Methodology

Because only a very small number of libraries and librarians are currently hosting text-digitization projects, the authors attempted to survey the entire population associated with publicly announced projects. To accomplish this, they identified survey participants from a number

TABLE 1
Job Titles of Study Respondents

Job Title or Major Responsibilities	#	%
Project coordinator	6	27.3%
Director/coordinator of electronic text center	4	18.2
Electronic text librarian	3	13.6
Academic computer center staff	2	9.1
Collection development librarian	2	9.1
Information technologies librarian	2	9.1
Preservation/conservation staff	2	9.1
Special collections librarian	1	4.5
Total	22	

of published and electronic sources.[31] The first round of the selection process resulted in forty candidates. From this population the authors eliminated six respondents for a variety of reasons such as their projects were not associated with academic libraries or they considered themselves more as "technical assistants" than librarians. Two participants with senior positions in libraries readdressed the survey to their colleagues, but these colleagues had been already included in the list and had responded. This resulted in thirty-two participants for the study.

Once the authors had identified the potential participants, they addressed the issue of survey format. After some thought, they opted to send an e-mail survey. When needed, the authors sent a second (or third) reminder by e-mail, two to three weeks after the initial survey. Three participants promised to answer questions and signed letters of consent but still did not respond; seven people did not answer at all. As a result, the authors had twenty-two answers to thirty-two surveys for a 68.8 percent return rate. Eleven out of twenty-two (50%) preferred that their comments be kept anonymous.

The twenty-two respondents fell into the work-related categories shown in table 1. Fourteen of the twenty-two participants (63.6%) had been involved with different kinds of digitization projects in the past, leaving eight (36.4%) who were totally new to this type of effort; nineteen out of twenty-two respondents (86.4%) were currently involved in various conversion projects, some working on several different projects at the same time; and three out of the twenty-two participants (13.6%) were planning to work on digitization projects in the future and had already planned their areas of focus for this work.

The survey covered a wide range of issues associated with the text-digitization process.[32] Questions focused on how materials are selected for digitization, how project teams convert works to an electronic form, which markup schemes and level of encoding (if any) are used, how materials are made accessible, and how materials are archived and maintained.

Findings and Discussion

Selection Issues

The process of digitization is logically preceded by the intellectually challenging procedure of selecting materials for such projects. Two survey questions specifically related to the difficulty of text selection: the first focused on what factors are involved in choosing a title; the second asked how respondents decide which editions of printed materials to digitize in these projects.

Respondents supplied the following information when asked to identify important selection factors (see table 2). They were asked to list all characteristics that were applicable to their decisions and to rank their responses.

The results show that the respondents thought the first three factors—intellectual coherence, collection strength, and library use—are almost equally important in selecting materials for digitization projects, with a slight emphasis on the "intellectual coherence" factor. The fol-

lowing comments show some of the different approaches possible in selecting materials:

> "at first we are interested in working with documents and collections that may have 'cachet' and 'high profile' so as to attract support, both moral and financial" (Palovitch, University of Pittsburgh);

> "We are largely demand driven in our approach, with some focus on collections of local strength. . . ." (Seaman, University of Virginia);

> "within a library context, especially in a library project to digitize a portion of its collection, collection strength is clearly very important" (name withheld).

In fortunate situations, a library may well be able to initialize its digitization efforts with materials that meet all three of these concerns: a collection may be splashy and thus "fundable" by trustees and alumni; it also may reflect serious scholarly use; and finally, it may represent a collection strength. Digitizing such materials will serve many purposes: fundraising, publicity for the library as a whole, outreach, scholarship, preservation of the physical material as the electronic version can be used in its place, and increased recognition of collecting strengths.

TABLE 2
Factors in Selecting Texts to Digitize (N=22)

Selection Factors	All Rankings #	All Rankings %	Top Rankings #	Top Rankings %
Intellectual coherence	17	77.3%	8	36.4%
Library's collection strength	17	77.3	7	31.8
Library use (the most frequently checked-out materials)	16	72.7	4	18.2
Other respondent-generated factors (e.g., "collection size," "long-term value," and "'sex appeal' of the item")	10	45.5	3	13.6

The survey also asked how the respondents selected specific editions (if applicable) to digitize. Again, they were to indicate all relevant answers (see table 3).

Although authority of the edition is clearly important to the respondents, copyright is frequently the key factor in deciding to digitize a work at all. Because "it is difficult and time-consuming to seek permission to republish printed works in electronic form on a text by text basis" (Seaman, University of Virginia), most respondents choose to work with texts out of copyright (mainly older documents available in the public domain) and to consider factors such as authority only secondarily. Many participants expressed their unhappiness and concern about this issue, and almost "envied" specialized projects that frequently have a legitimate, scholastic reason to work with materials in the public domain. For example, for the University of

TABLE 3
Factors in Selecting Editions to Digitize (N=22)

Selection Factors for Editions	#	%
Most authoritative	15	68.2%
Anything out of copyright	13	59.1
Most recent edition	7	31.8
Other respondent-generated factors (e.g., "most intellectually interesting," "the original editions," and "whatever is available")	7	31.8

TABLE 4
Approaches to Text Conversion (N=22)

Approach to Capturing Text	#	%
Combination of OCR and Bit-map Scanning	15	68.2%
OCR	10	45.5
Bit-map Scanning	9	40.9
Other respondent-generated factors (e.g., "keyboarding")	7	31.8

North Carolina (UNC-CH) pilot e-text project, the decision was made to use only "texts printed during the nineteenth century that were actually read by people at the time and influenced their thinking" (Dominguez, UNC-CH). Numerous written comments showed that until the copyright issues are cleared, most project managers will remain "unhappy about selecting out-of-copyright materials" unless their users' needs demand digitization of these texts. As one respondent noted, "we need good scholarly resources and those are not often the ones that are out of copyright."

If keying is to be done, will it be an in-house project, outsourced, or sent offshore for less expensive labor rates?

An extended discussion of copyright is beyond the scope of this article, but it may well be the largest problem associated with e-text projects. This is an area in which librarians and scholars must work with publishers in order to make high-quality materials available in an electronic format, whether the publishers or a third party, such as a library, encodes and disseminates them.

Conversion Issues

Text-conversion decisions are the core issues of creating e-text collections in any setting, including libraries. Before a digitization process starts, the project coordinator should answer many important questions, such as: How will the text be converted into computer-readable form? Will the text be hand-keyed or scanned? If scanned, will it be scanned with optical character recognition (OCR) software, or will it be scanned as bit-map images that are not searchable text? How will a high level of accuracy be achieved? If keying is to be done, will it be an in-house project, outsourced, or sent offshore for less expensive labor rates?

One question in the survey dealt with different approaches to capturing text and converting it into computer-readable form. Participants were asked to check as many options as they had used or would consider using in the future. The results are presented in table 4, with the most frequently mentioned first.

The majority of respondents suggested the combination of OCR and bit-map images as the best way to prepare electronic texts. Presenting collections in both forms—as searchable ASCII text from OCR conversion (and eventually encoded text) and as digital bit-mapped page images—has been implemented in various projects at the University of Virginia (UVa), and is the goal of the UNC-CH pilot project on digitizing nineteenth-century southern Americana.

Many practitioners noted that people are excited about OCR until they start using it. Specifically, OCR accuracy is only reliable if one starts with perfectly clear and clean-printed documents. This is not the case, however, in the majority of digitization projects that deal mainly with older typeset material. With nineteenth-century books, the results may vary from 80 to 95 percent accuracy, which in the best case represents a few misspellings per line—and this is after training the software program to recognize the particular text font used! The frustration with this generally inadequate level of accuracy for automated conversion processes leads many project coor-

dinators to decide in favor of keyboarding, using either in-house staff or hiring contractors, here in the United States or offshore. To reach higher accuracy levels, some projects use double keying, or a combination of OCR and keyboarding (as is the case with some projects at UVa). It is worth mentioning that along with text-conversion problems go the issues of acceptable accuracy and the tremendously time-consuming editing and quality control processes.

Data Presentation Issues

Data presentation, the format in which e-texts exist, is one of the most important considerations for a high-quality text. The vast majority of e-texts today are available only in ASCII format. Susan Hockey estimates that "the remaining 5 percent of texts are what can be called packaged products, where the text has been indexed for use with specific, often proprietary, software and cannot be used for any other purpose."[33] Although ASCII text is portable from one system or computer to another, most formatting information is lost in the conversion from either nonelectronic hard copy or most proprietary systems to markup-free ASCII text; thus, this is not a perfect solution for humanities texts that contain many non-ASCII characters and are rich in typographic features. Conversely, nonportable, proprietary or idiosyncratic software, encoding, or markup prohibit the exchange of data.[34] Anyone planning an e-text project needs to address several text presentation issues before beginning, including what markup scheme (if any) should be applied and what level of encoding is appropriate. The survey used three questions to elicit respondents' opinions regarding the very complicated issues associated with text presentation.

Markup and SGML

Over the past forty years, scholars have struggled with a wide variety of markup formats, very often each designed for a specific project. As a general rule, e-texts have been poorly documented, and users have wasted their time trying to convert texts from one format to another. Finally, in 1986, after almost twenty years of work, SGML (Standard Generalized Markup Language) became a standard of the International Organization for Standardization (ISO 8879).[35]

What is SGML, and why is it so special? Three characteristics distinguish SGML from any other markup language today:

1. its emphasis on descriptive rather than procedural markup;
2. its document type concept; and
3. its independence of any one system for representing the script in which a text is written.[36]

Over the past forty years, scholars have struggled with a wide variety of markup formats, very often each designed for a specific project.

SGML accommodates numerous "document-type definitions" (DTDs) so that various types of documents can be made to appear just as they might in hard copy with all formatting intact. Because of its extremely complex set of rules, "SGML is potentially the most powerful publishing technology around," but it also has been "the least understood, least appreciated, and least sexy."[37] At the same time, the object-oriented, rule-based, open, and very logical system can be used to represent virtually any data we may encounter today or in the future. The fact that "a number of companies, including Microsoft and Novell/Word Perfect, are or soon will be providing products with SGML support"[38] testifies to its future as an encoding "metalanguage." For more information about SGML and related topics, visit the International SGML User's Group.[39] Various discussion groups and listservs (such as SGML Newswire

TABLE 5
Preferred Encoding Schemes (N=22)

Encoding Schemes Used	#	%
SGML	18	81.8%
HTML	11	50.0
Other	3	13.6
None	2	9.1

and SGML-L) are concerned with relevant topics, and there is a growing literature discussing SGML.[40]

Responses to the first question about encoding schemes the respondents had used or would use to mark up large amounts of text are presented in table 5. Interestingly, one of the respondents considered a new emerging European standard, Open Document Architecture (ODA), as the most suitable for his U.S.-located project.

The results demonstrate that SGML, ISO 8879, is the leading encoding scheme for e-text markup today. It is rightfully gaining authority and rapidly becoming the standard for managing, distributing, and publishing documents. Readers should be cautioned that HTML (HyperText Markup Language) ranks as high as it does only because many respondents considered it as a delivery medium, and at the time of the survey no SGML browsers for the World Wide Web existed. Since then, SoftQuad marketed the first SGML browser, Panorama, and opinions on this question may well have changed.[41] HTML is only one of many SGML applications, one of its DTDs, and quite a simplified one. Recently, many specialists have expressed their hope that a scaleable HTML will be able to accommodate and present complex Web documents.

TABLE 6
Typical Level of Text Encoding (N=22)

Level of Encoding	#	%
Minimal level of detail	9	40.9%
In-depth level of detail	8	36.4
Other	6	27.3
Undecided	1	4.5

Levels of Encoding

The second data presentation question, concerning levels of encoding implemented in digitization projects, evoked a general concern regarding costs. Participants also expressed an almost unanimous opinion that the selection of an encoding level depends on the nature of a project and user requirements. Except to meet specific research needs, many respondents said that a minimal level of detail was the only affordable approach. The results to this question are summarized in table 6.

As well as limiting processing costs, encoding at a minimal also can result in more tailored texts. Scholars can take a text that has minimal encoding and subsequently add additional tags to meet their specific research needs without cost to the overall project. Of course, using a standardized markup scheme is central to the widespread effectiveness of this approach.

The Text Encoding Initiative

The Text Encoding Initiative (TEI), an international cooperative project, was established in 1987.[42] A particular advantage of the project is that scholars conceived it and started it by themselves. The importance of this fact cannot be overestimated. Because the initiative is based on "inside" knowledge of humanists' needs, it should serve researchers, as well as teachers and students in the humanities. The decision of the TEI editors to use the SGML international standard gives the project a special strength and needed authority.

The "TEI Guidelines" are built on the principle of descriptive markup, with the goal that any properly encoded text should be transportable from one hard-

ware or software environment to another without losing any information. The idea of so-called durability or reusability of e-texts is extremely important to a rapidly growing audience. This crucial concept means that texts are prepared in such a manner that they "will last a long time and that [they] can be used for as many applications as possible, which would lead to issues of improving intellectual access."[43]

The guidelines provide recommendations for encoding a rich variety of literary forms and genres—not only prose texts, but also verse, drama, and other performance texts, transcripts of spoken material for linguistic research, dictionaries, and large corpora. The creators/encoders of an e-text are responsible for selecting, from among the tag sets, those that are applicable to specific documents and the specific goals of an encoding project. Charles Goldfarb captures the importance of the "TEI Guidelines" for cultural information:

> The vaunted "information superhighway" would hardly be worth traveling if the landscape were dominated by industrial parks, office buildings, and shopping malls. Thanks to the Text Encoding Initiative, there will be museums, libraries, theaters, and universities as well.[44]

The TEI editors state that the guidelines "answer the fundamental needs of a wide range [of] sciences, publishers, librarians, and those concerned generally with document retrieval and storage."[45] More than four hundred tags give a detailed specification for the documentation of practically any type of electronic materials, their sources, and their encoding—information unfortunately neglected by the authors of numerous other encoding metalanguages. Librarians will particularly appreciate specifications for provenance of e-texts in this encoding scheme because they provide necessary information for developing reliable and efficient cataloging procedures. (This topic is discussed in the following section.)

The next question asked respondents how they would present e-texts. Respondents were encouraged to check all options they would consider applicable. Here, fifteen respondents (68.2%) indicated that they would make texts conformant with "TEI Guidelines"; eight (36.4%) said they would use some other presentation format; and one was unsure of a choice for this question. The intention of the question was mainly to test the adoption rate of the "TEI Guidelines" in academic library settings. The results show the direction chosen by the majority of participants in the survey, i.e., create e-text collections conformant with "TEI Guidelines."

Access Issues

The constantly growing number of electronic texts (and electronic resources in general) forces individuals and institutions, and especially libraries, to find effective ways and tools to control the avalanche of information. Currently, professionals are experimenting with a variety of approaches to this problem. An ALA committee, MARBI (Machine-readable Bibliographic Information), advises the Library of Congress on additions and changes to the USMARC format and, in 1993, the "Electronic Location and Access" 856 field was included in the USMARC format for Internet-accessible materials. Various professional meetings, some sponsored by the Library of Congress, help librarians, especially catalogers, approach difficult e-text issues.[46]

In 1994, the U.S. Department of Education awarded a $62,000 College Library Technology and Cooperation grant to support the OCLC project, Building a Catalog of Internet Resources.[47] The aim of the project is to create, test, and later implement a searchable database of USMARC format bibliographic records

TABLE 7
Preferred Methods of Providing Bibliographic Access to Library-Created E-texts (N=22)

Bibliographic Access Method	#	%
Local OPAC	9	40.9%
OCLC	9	40.9
Undecided	5	22.7
RLIN	3	13.6
Internet	3	13.6
Stand-alone	1	4.5
Other	1	4.5

containing information on electronic location and access to resources available over the Internet. CETH has undertaken other activities important to providing access to electronic materials. It has developed procedures for cataloging e-texts in the MARC format and placing records in the RLIN (Research Libraries Information Network) union catalog.[48]

Various efforts to create catalogs for electronic materials include *Alex: A Catalog of Electronic Texts on the Internet*, consisting of almost 1,800 entries, begun in 1994.[49] Unfortunately, maintenance of the Alex catalog has been suspended until funding can be obtained. Other interesting attempts to catalog electronic materials include CATRIONA (CATaloguing and Retrieval of Information Over Networks Application) from the British Library Research and Development Division and the Electronic Resources Project at the University of Toronto.[50] Columbia University Library also is involved in a project to catalog e-texts.[51]

Even a quick look at available catalogs leads to the issue of what documentary information should be included for providing consistent bibliographic control. Is MARC the way to go? What changes are needed? Invaluable help comes from the "TEI Guidelines," and specifically from the so-called TEI header that addresses "the problems of describing an encoded work so that the text itself, its source, its encoding, and its revisions are thoroughly documented."[52] The "TEI Guidelines" go on to note that "such documentation is equally necessary for scholars using the texts, for software processing them, and for cataloguers in libraries and archives. Together these descriptions provide an electronic analogue to the title page attached to a printed work."[53] In order to be TEI conformant, every e-text must have a TEI header.[54]

Recently, librarians have had some heated discussions over the relationship between the "TEI Guidelines" and MARC. A few projects are attempting to apply the principles of the TEI and USMARC to provide consistent bibliographic control for e-text collections. These examples include the British National Corpus, the CETH, and the University of Virginia cataloging procedures.[55]

The survey sought respondents' opinions regarding possible ways to approach the problem of providing bibliographic access to library-created e-texts. Participants were asked to check as many options as they would find applicable. The results are presented in table 7, in descending order.

Archiving Issues

With the inevitable "migration" of e-texts from one site to another, archiving issues become increasingly important. Al-

TABLE 8
Preferred Archive Site for Library-Created E-texts (N=22)

Archive Site	All Rankings #	All Rankings %	Top Rankings #	Top Rankings %
Library/producer/ owner	18	81.8%	16	72.7%
Computer center	9	40.9	3	13.6
Other	2	9.1	3	13.6

though electronic publishing in libraries is too recent a development to make any long-term predictions, the authors have tried to identify state-of-the-art facilities. Today, there are two main sites able to carry responsibility for archiving library-created e-texts—libraries and computer centers. Historically, libraries have played the role of repositories and archives of human knowledge, and have used a scientific approach to organizing materials. At this time, we are talking about a new type of information, and presently, computer centers may provide better equipment and professional expertise for archiving electronic resources than libraries can. This delicate situation should lead to a closer relationship between the two players, one that is beneficial for both sides. In the future, as one of the participants in the survey pointed out, we may witness "this role gradually moving to the library as libraries acquire more computer expertise." Others were even more explicit: "I feel very strongly that this is a library's task in an electronic world" (N. Finke, director for electronic text in the law, University of Cincinnati College of Law).

The survey sought respondents' views regarding who should be responsible for archiving library-created e-texts. The results are presented in table 8, with the most frequently mentioned option first.

Along with the issue of archiving library-created e-texts arises the issue of documenting upcoming revisions of the files. With the relative ease of making changes in electronic documents, who has responsibility for updating and refreshing the data? The revision description should carefully document the revision history for the file, and constitutes one of the four core parts of the so-called TEI header, which gives important information for librarians, users, and computer scientists. The information on corrections and changes should be provided and documented by producers and everyone participating in the process of a text's revision, whether a scholar, a librarian, or anyone else.

The survey asked who has responsibility for a text's revision history. The results, organized in descending order, are shown in table 9.

Staff and User Education Issues

Staff training has become one of the most urgent issues since computers were first installed in libraries. Two issues—whom to train and how—need to be addressed immediately. Relevant discussions about staffing have taken place in conferences and professional meetings since the first e-text centers opened their doors to the public. In most academic libraries, electronic information services have been set up in reference departments, and reference librarians have been the first group to learn to use these tools and resources. Reference librarians have had to learn many new skills because they have been expected to execute new and unusual operations, including: how to use these databases, how to create data files, what

Recently, librarians have had some heated discussions over the relationship between the "TEI Guidelines" and MARC.

TABLE 9
Organization Responsible for Revising E-texts (N=22)

Responsible for Text's Revision History	All Rankings #	All Rankings %	Top Rankings #	Top Rankings %
Library/producer/ owner	20	90.9%	15	68.2%
Computer center	7	31.8	5	22.7
Other	2	9.1	1	4.5
Undecided	1	4.5	1	4.5

advice to give on different interfaces, how to help with markup systems and available software programs, and even how to provide "frontline" technical support. Librarians will be able to carry out these operations more effectively through closer cooperation with the library systems offices and the university or college computing centers.

Recently, more and more opportunities have arisen for staff training/education regarding e-texts. One of the most important is the intensive two-week seminar offered every summer, since 1992, at the CETH facilities. Its main aims are "the development of educational programs to provide support for librarians, scholars, and teachers who are working with electronic texts."[56] The seminar "addresses a wide range of challenges and opportunities that electronic texts and software offer [and] covers such areas as data capture, markup, retrieval, presentation, transformation, and analysis of electronic text."[57] The Rare Books School at UVa also offers a program on e-texts. Attending conferences sponsored by the Association for Computers and the Humanities[58] and other professional meetings, participating in electronic-specialized discussion groups and listservs, and networking with colleagues at other academic libraries will give librarians a better vision and understanding of working with e-texts.

TABLE 10
Preferred Methods of Staff and User Instruction

User Instruction Methods	#	%
In the library at point of use	19	83.4%
Workshops	15	68.2
Computer-assisted instruction	10	45.4
In academic departments	9	40.9
Other	4	18.2

Another solution may come from professional schools educating new generations of librarians. New roles, functions, and even more—new concepts of librarianship—are forcing academic programs to reshape their curriculums to accommodate material concerning e-texts and access to them. It is a very important task for information and library science schools to undertake—and a difficult one.

Historically, the user education mission has been associated with libraries. With the presence of new technology in libraries, this mission becomes more urgent as staff from UVa's e-text center note:

> Through ongoing training sessions and support of individual teaching and research projects, the Center is building a diverse and expanding user community locally, and providing a potential model for similar enterprises at other institutions.[59]

The survey asked respondents to check the best ways to offer staff and user education (see table 10). They were encouraged to check as many options as they would recommend. In addition, some participants made very interesting suggestions such as: using online documentation, using written documentation supplied with the data, and offering self instruction on a Web server.

Other Issues to Consider

Many respondents suggested additional challenges that librarians need to consider in the process of preparing e-texts. In summary, their hopes focus on solving copyright problems, protecting intellectual property, providing file security, improving software programs that are still very primitive in comparison with the intellectual needs of scholars, using new imaging technologies, preserving files, distributing information, coordinating so as to avoid duplication of effort, among many others. The most urgent comments addressed the need to set up standards and quality control.

Conclusions

We live in an exciting and challenging time of continuously developing information technology that affects our lives and daily activities. Tremendous changes are occurring on academic campuses that influence scholars' communication patterns and the entire scholarly research process. Despite the widespread opinion that humanists are reluctant to use technology in general, we see how the humanities disciplines are becoming increasingly computer dependent. Online services, CD-ROMs, hypermedia, bulletin boards, and local, national, and international networks are changing the way humanists do research and teach. Indeed, these technologies change the way they use information. One of the areas experiencing the most change is that of e-texts, both in the sense of how scholars are using them and in how they are being produced.

The proceedings of the conference titled "Technology, Scholarship, and the Humanities: The Implications of Electronic Information" state that one of the main themes common to all five working groups was to "promote, as a national priority, the creation of a 10-million-volume digital library, broadly conceived to encompass the full spectrum of humanities research collections."[60] This has proven to be too large a goal for at least the near future, but progress is being made. Numerous technical, legal, and financial obstacles will need to be overcome, however, before we will see the 10-million-volume digital library on our virtual shelves. We are still in the earliest stage of building any universal electronic library. However, the process has started and is irreversible. The initiation of this process raises many issues that libraries must address if they are to continue to play a vital role in serving humanists.

First, with the promise of an electronic, paperless world, libraries must rethink their traditional role of being only "archival repositories" of information and decide whether they are to become producers of e-texts as well. The electronic revolution may provide libraries with an opportunity to take a "business" approach, analogous to that of commercial publishers. The technological evolution on campuses, together with a three-way partnership among the library, the computing center, and the faculty, offers new potential for creating university publishing enterprises. However, this sort of enterprise will become a reality only if these three groups join forces.

One of the areas experiencing the most change is that of e-texts. . . .

Once libraries make the decision to create e-texts in-house to meet their clients' needs as well as to archive such materials, and when the necessary linkages and infrastructure are developed, several specific issues must be addressed. First, project directors must decide which printed texts to convert to an electronic format. To make this a more rational process, programs should have a text creation policy (part of a larger e-text collection policy) that sets forth various guidelines. Respondents to the authors' survey indicated the importance of matching e-text collection focus with overall collection strength and meeting the needs of local library users, although such materials may be made available worldwide. Within these considerations, intellectual coherence of the work itself was seen as most important. Once a title is selected for digitization, the respondents noted that they would seek the most authoritative, scholarly work to include in their e-text collections; but on a practical level, they would most likely have to select materials that were out of copyright. Copyright law appears to be a major roadblock hindering the digitization of library materials.

Along with collection considerations, project directors also must decide how

to convert texts to an electronic format. The three most compelling factors here are cost, accuracy, and searchability. Bit-mapped images are the least expensive and most reliable in terms of information/image content, but they do not yield text that is searchable and there is no widely accepted international standard for bit-mapped images. Searchable text requires either human keyboard entry or optical character recognition processing. Both processes are expensive and require extensive checking in the form of double keying or scanning or manual proofreading in order to attain the high degree of accuracy most humanists require in their e-texts. Until OCR software is much improved, high-quality text conversion will remain a very expensive proposition and one that greatly limits the extent of e-text programs and their collections.

The ability of a project to conduct in-depth coding undoubtedly relates to its financial support because this can be a very costly undertaking.

Once text has been converted to electronic form, projects must decide how to encode the information. Most current projects advocate the use of SGML coding and, to a lesser degree, HTML. Since the time of the survey, the appearance of an SGML browser for the Web (Panorama from SoftQuad) has led more projects to use SGML because it retains much more of the visual richness of a document than HTML does and provides full-text search capabilities within entire documents and specific SGML elements, such as titles, tables of contents, notes, or chapters. Along with a preferred encoding scheme, projects also must select an appropriate level of text encoding. The respondents were split almost in half over this question, some indicating they typically used a minimal level and an almost equal number advocating in-depth encoding. The ability of a project to conduct in-depth coding undoubtedly relates to its financial support because this can be a very costly undertaking. Of course, if everyone is using SGML as the standard for their initiatives, scholars can possibly take minimally encoded text and add additional coding as required by their work. This clearly indicates the importance of projects adhering to the SGML international standard and specifically, the "TEI Guidelines," as the application of SGML for texts in the humanities.

The projects mentioned throughout this paper demonstrate that academic libraries are already in the business of creating e-texts. Two developments now offer invaluable help in this new endeavor for libraries. First, CETH serves as a major coordinator of international efforts in compiling, maintaining, and disseminating e-texts. Second, TEI, an international cooperative project, helps with applying standards and quality control in the process of creating e-texts.

Billington reminds librarians of their responsibility as leaders in this networked era: "Through the rich resource that libraries represent, the information superhighway can give a new boost, born of access to knowledge, that will feed the intellectual curiosity, entrepreneurial energy and civic spirit of Americans in the 21st century."[61] This is our duty.

The authors would like to thank Pat Dominguez, Susan Hockey, Michael Neuman, Sydney Pierce, and Michael Sperberg-McQueen for their thoughtful reading of this text and their insightful comments.

Notes*

1. Clifford Lynch, "Networked Information: A Revolution in Progress," in *Networks, Open Access, and Virtual Libraries: Implications for the Research Library*, eds. Brett Sutton and Charles H. Davis (Urbana, Ill.: Graduate School of Library and Information Science, 1992), 12–13.

2. James H. Billington, "The Electronic Library," *Media Studies Journal* 8 (winter 1994): 109.

3. ———, "Electronic Access: Dr. Billington Testifies on the Role of Libraries," *LC Information Bulletin* 53 (May 16, 1994): 187.

4. *Humanities and Arts on the Information Highway: A Profile, Final Report*, a national initiative, sponsored by the Getty Art History Information Program, the American Council of Learned Societies, the Coalition for Networked Information, September 1994 (Santa Monica, Calif.: J. Paul Getty Trust, 1994), 2; see also, URL: gopher://gopher.cni.org:70/00/cniftp/projects/humartiway/humartiway-rpt.

5. See the following URLs for periodic reports on these programs: http://www.grainger.uiuc.edu/dli/; http://elib.cs.berkeley.edu/; http://http2.sils.umich.edu/UMDL/; http://fuzine.mt.cs.cmu.edu/im/informedia.html; http://alexandria.sdc.ucsb.edu; http://diglib.stanford.edu/diglib/.

6. "Bell Atlantic Gives $1.5 Million to NDL Program," URL: http://lcweb.loc.gov/ndl/oct-95.html#bell.

7. Helen R. Tibbo, "Information Systems, Services, and Technology for the Humanities," *Annual Review of Information Science and Technology* 26 (Medford, N.J.: Learned Information, 1991): 290.

8. *Computers and the Humanities* (New York: Pergamon, 1966).

9. "Literary and Linguistic Computing," in *Journal of the Association for Literary and Linguistic Computing* (Oxford: Oxford Univ. Pr., 1986).

10. Humanist listserv: humanist@lists.princeton.edu. See also, the Humanist Web site: URL: http://www.princeton.edu/~mccarty/humanist/.

11. URL: http://www.ceth.rutgers.edu/info/ectrdir.html.

12. Katharina Klemperer, "Electronic Texts: Introduction," *Information Technology and Libraries* 13, no. 1 (Mar. 1994): 6.

13. URL: http://www.unc.edu/~nsmith/.

14. Dan Ream, "The University of Virginia's Electronic Text Center: An Interview with David Seaman," *Virginia Librarian* 39 (Apr.–June 1993): 7.

15. David M. Seaman, "The Electronic Text Center: A Humanities Computing Initiative at the University of Virginia," *Electronic Library* 11 (June 1993): 195–99; and "'A Library and Apparatus of Every Kind': The Electronic Text Center at the University of Virginia," *Information Technology and Libraries* 13 (Mar. 1994): 15–19. See also, URL: http://www.lib.virginia.edu/etext/ETC.html.

16. The Humanities Text Initiative, begun at the University of Michigan (UM) in the fall of 1994, seeks to bring textual analysis resources to the UM campus, "focusing on networked delivery of text and text encoding in a standardized format." URL: http://www.hti.umich.edu/hti/vitality.html. See also Beth Forrest Warner and David Barber, "Building the Digital Library: The University of Michigan's UMLibText Project," *Information Technology and Libraries* 13 (Mar. 1994): 20–24.

17. Michael Neuman, "The Very Pulse of the Machine: Three Trends toward Improvement in the Development of Electronic Versions of Humanities Texts," *Computers and the Humanities* 25 (spring 1992): 363–75; Michael Neuman and Paul Mangiafico, "Providing and Accessing Information via the Internet: The Georgetown Catalogue of Projects in Electronic Text," *Reference Librarian* 41/42 (1994): 319–32. See also, URL: http://www.georgetown.edu/guhome.html.

18. Indiana University's Victorian Women Writers Project can be found at: http://www.indiana.edu/~letrs/vwwp/.

19. See CETH electronic distribution list (list owner: Pamela Cohen, Center for Electronic Texts in the Humanities, New Brunswick, N.J.) and *CETH Newsletter* (New Brunswick, N.J.: Center for Electronic Texts in the Humanities, spring 1993–).

20. Marianne I. Gaunt, "Center for Electronic Texts in the Humanities," *Information Technology and Libraries* 13 (Mar. 1994): 9.

21. Susan Hockey, "Evaluating Electronic Texts in the Humanities," *Library Trends* 42, no. 4 (spring 1994): 677.

22. Gaunt, "Center for Electronic Texts," 7.

23. Oxford Text Archive, *A Shortlist of Machine-Readable Texts Held at Oxford Text Archive* (Oxford: Oxford Text Archive, Oxford Computing Service, 1994). See also, URL for ftp:// ota.ox.ac.uk.

24. Susan Hockey, "Developing Access to Electronic Texts in the Humanities," *Computers in Libraries* 13 (Feb. 1993): 41–43.

25. URL: http://users.ox.ac.uk/~archive/ota.html.

26. Ibid.

27. John Price-Wilkin, *The Feasibility of Wide-area Textual Analysis Systems in Libraries: A Practical Analysis*, URL: http://www.lib.virginia.edu/staffpubs/jpw/dpc.html.

28. Judith Zidar, "Text Conversion," in *Workshop on Electronic Texts: Proceedings*, ed. James Daly (Washington, D.C.: Library of Congress, 1992), 65. See also, URL: gopher://marvel.loc.gov:70/00/loc/pubs/am/e-text/e-sess4.txt.

29. Ibid.

30. Guy Lamolinara, "Metamorphosis of a National Treasure," *American Libraries* 27 (Mar. 1996): 32.

31. The following sources were used to identify participants for the survey: *Workshop on Electronic Texts. Library of Congress. Proceedings. Appendix III: Directory of Participants*, June 9–10, 1992; *The Georgetown Catalogue of Projects in Electronic Text (CPET): Electronic Text Projects in Literature* (information on the projects with contact address, etc.); *The CETH Catalog of Electronic Text Centers in the U.S.*; descriptions of relevant projects in specialized literature; and introductions of relevant projects in professional discussion groups such as ETEXT.

32. Due to space limitations, the survey was not included in this article. Please contact the authors for a copy.

33. Hockey, "Evaluating Electronic Texts," 678.

34. *Encoding*, or *markup*, is the explicit formatting of information contained in a computer file other than the text itself. These instructions tell an output device, such as a screen or printer, how the material should appear.

35. ISO 8879, *Information Processing-Text and Office Systems-Standard Generalized Markup Language (SGML)* (Geneva: International Organization for Standardization, 1986). See also, Sharon Adler, "The Birth of a Standard," *Journal of the American Society for Information Science* 43 (Sept. 1992): 556–58.

36. C. M. Sperberg-McQueen and Lou Burnard, eds., *Guidelines for Electronic Text Encoding and Interchange* (Chicago and Oxford: Text Encoding Initiative, 1994), 14.

37. Lori Grunin, "Publish without Paper," *PC Magazine* 14 (Feb. 7, 1995): 112.

38. Ibid., 116.

39. URL: http://www.sil.org/sgml/sgml.html.

40. See, for example, Martin Bryan, *SGML: An Author's Guide to the Standard Generalized Markup Language* (Wokingham, England: Addison-Wesley, 1992); Charles F. Goldfarb, *The SGML Handbook* (Oxford: Clarendon Pr., 1993); Anneli Heimbürger, "Introduction to Standard Generalized Markup Language (SGML)," *Microcomputers for Information Management* 11/4 (1994): 239–60; Eric van Herwijnen, *Practical SGML*, 2nd ed. (Dordrecht: Kluwer Academic Pubs., 1994).

41. URL: http://www.sq.com/products/panorama/panor-fe.htm.

42. See Malcolm B. Brown, "What Is the TEI?" *Information Technology and Libraries* 13 (Mar. 1994): 8.

43. Susan Hockey, "Approaches to Preparing Electronic Texts," in *Workshop on Electronic Texts: Proceedings*, ed. James Daly (Washington, D.C.: Library of Congress, 1992), 66. See also, URL: gopher://marvel.loc.gov:70/00/loc/pubs/am/e-text/e-sess5.html.

44. Charles F. Goldfarb, "Preface," *Computers in the Humanities* 29/1 (1995): 1.

45. URL: http://www-tei.uic.edu/orgs/tei/info/guide.html.

46. URL: http://lcweb.loc.gov/catdir/semdigdocs/seminar.html.

47. More information on this project can be found through the OCLC Web site at http://www.oclc.org/oclc/man/catproj/catcall.htm.

48. The *Guidelines for Cataloging Monographic Electronic Texts* are available in PostScript version at http://cethmac.princeton.edu/Docs/catguid.ps; or as an ftp file (ftp to ceth.princeton.edu). For the print version, see Annelies Hoogcarspel, *Guidelines for Cataloging Monographic Electronic Texts at the Center for Electronic Texts in the Humanities*, CETH Technical Report no. 1 (New Brunswick, N.J.: CETH, 1994).

49. URL: http://www.lib.ncsu.edu/stacks/alex-index.html.

50. CATRIONA URL: http://www.bubl.bath.ac.uk/BUBL/catriona.html; ERP: http://www.fis.utoronto.ca/ejournals/.

51. "Columbia Catalogs Electronic Texts," *Library Journal* 119 (Feb. 1, 1994): 22.

52. Sperberg-McQueen and Burnard, eds., *Guidelines for Electronic Text Encoding and Interchange*, 89.

53. Ibid.

54. Richard Giordano, "The Documentation of Electronic Texts Using Text Encoding Initiative Headers: An Introduction," *Library Resources and Technical Services* 38 (Oct. 1994): 389–401.

55. Edward Gaynor, "Cataloging Electronic Texts: The University of Virginia Library Experience," *Library Resources and Technical Services* 38 (Oct. 1994): 403–13; Annelies Hoogcarspel, "The Rutgers Inventory of Machine-Readable Texts in the Humanities: Cataloging and Access," *Information Technology and Libraries* 13 (Mar. 1994): 27–34. See also "Berkeley Finding Aids Conference," URL: http://sunsite.berkeley.edu/FindingAids/.

56. Gaunt, "Center for Electronic Texts," 11.

57. Ibid.

58. URL: http://www.ach.org.

59. URL: http://www.lib.virginia.edu/etext/ETC.html.

60. "Themes Common to the Five Working Groups," in *Technology, Scholarship, and the Humanities: The Implications of Electronic Information*. Conference, Sept. 30–Oct. 2, 1992. Summary of Proceedings (Santa Monica, Calif.: The American Council of Learned Societies and the J. Paul Getty Trust, 1993), 23.

61. Billington, "Electronic Access," 112.

*All cited URLs were valid as of August 8, 1996.