CATALOGING FOR DIGITAL LIBRARIES:
THE TEI SCHEME AND THE TEI HEADER

LINE POUCHARD
UNIVERSITY OF TENNESSEE, KNOXVILLE

This article describes the uses and advantages of using the Text Encoding Initiative (TEI) guidelines for cataloging electronic texts. The TEI guidelines have been developed through an international and collaborative effort, and their applications in digital libraries such as the University of Virginia Electronic Text Center have required close collaboration between catalogers and humanities computing researchers. Detailed description and examples of the TEI header, a vehicle for meta-information written in SGML and the part of the TEI scheme most useful to librarians, are provided. Possible congruence between TEI headers and USMARC records implies that granularity of the TEI header and flexibility of the MARC record are simultaneously improved.

INTRODUCTION

Every document made available for the purpose of library collection, in electronic form or otherwise, must satisfy requirements of stability, source reliability, and bibliographic information in order to be useful to a community of users, not just to the person who created the document. Scholars in the humanities have long been interested in the possibilities of storing and retrieving textual materials in electronic form but many of the electronic documents originally encoded did not satisfy these requirements. In particular, digitizing efforts of rare material, old manuscripts, and early editions, for the most part, were not useful to scholars because these electronic documents contained many spelling errors (as a result of the image-to-text conversion process), omitted the publication information, and the edition of the printed text that was used as a basis for the electronic document.

One use of electronic texts in the humanities is the possibility of automatically compiling textual variants for new editions of ancient works, and another is making accessible to scholars material which, due to the location of the item, would make it difficult for them to consult physically. It is therefore of prime importance that the electronic materials be accurate in their transcriptions, and contain detailed and accurate bibliographic information. The bibliographic information required by scholars of ancient texts also needs to be more descriptive than a bibliographic citation for library purposes.

In addition to traditional bibliographic information such as that provided by proper cataloging of an electronic text, humanities scholars need more detailed description of the physical book or manuscript they study. For instance, information such as where a page break occurs in the original editions and manuscripts of the same work, the appearance of the title page, and how the lines of verse or prose are arranged on the page in a play or poem, are only a few items indispensable to a scholar who tries to compile a new edition of an ancient work. While this type of meta-information is not in the scope of cataloging per se, it is required, for instance, for an electronic version of a first quarto of a Shakespearean play to be useful. Electronic texts for the humanities thus create problems as to how much and what type of meta-information they must provide.

In the context of this specific need for accurate and detailed meta-information to accompany electronic texts, the Text Encoding Initiative (TEI) was established in 1987 as a volunteer effort in humanities computing, an effort driven by a joint-project of the Association for Computers and the Humanities, the Association for Computational Linguistics, and the Association for Literary and Linguistic Computing. The goal of the TEI project is "to define a set of generic Guidelines for the representation of textual materials in electronic form" (Burnard, 1995, part 2).

This article discusses some elements of a document encoded with TEI that are appropriate for bibliographic description (i.e., the TEI header) and presents an example of a TEI-encoded document. It also discusses the scheme for meta-information included in the TEI header and compares the encoding of the document according to the Text Encoding Initiative with a record established for the same document using USMARC.

SGML AND THE
TEXT ENCODING INITIATIVE

The TEI guidelines specify a number of required elements for each text, in particular a header made of four required elements which contain the bibliographic information and a body which can contain text, images, and other objects. In order to provide rules for the encoding and interchange of electronic texts, the TEI scheme relies upon the use of SGML (Standard Generalized Mark-up Language) and its sets of mark-up tags for the encoding of textual material. SGML is an international standard (ISO 8879) for encoding electronic information which defines device- and systems-independent methods for representing text and other objects and is concerned with content and arrangement rather than format or appearance. SGML is made of various subsets of tags, called Document-Type Definitions (DTD) which specify content requirements and sets of tags for each type of document. HTML (Hypertext Mark-up Language), which is one type of an SGML DTD, is the best-known sub-set of SGML tags. The TEI guidelines are another.

Like other SGML applications, TEI is independent of platforms, systems, applications and devices and conforms to network protocols for the exchange of information. Any electronic document encoded in SGML includes 3 parts: the SGML declaration (i.e. a statement declaring this is an SGML document), a DTD (i.e., the sets of SGML tags used for this particular document), and the document instance (i.e., the encoded document according to the rules declared by the DTD). With SGML, and therefore with TEI, each electronic document carries along its own meta-information or metadata.

The TEI scheme defines how to write a specific class of SGML document-type-definitions that specify how each electronic document is encoded in SGML. According to the TEI guidelines, the content requirements specify minimum mark-up requirements for a low level of encoding, but may also provide for very complex encoding appropriate to the detailed marking necessary to

humanities research (Horowitz & Palowitch, 1996). Each electronic document encoded according to TEI also contains a TEI header in addition to the body of the document. TEI headers carry meta-information about the electronic document, contain bibliographic information that is of direct use to libraries, and has been designed in consultation with librarians.

Although the TEI guidelines were originally concerned with printed texts such as those studied in literature, linguistics, and history, and designed as "a common encoding scheme for complex textual structures" (Burnard & Sperberg-McQueen, 1994, preface), electronic texts must be understood in a broad way: TEI is interested in textual and non-textual resources such as those contained in a research database or components of non-paper publications (Burnard, 1995, part 2).

As the project developed, more and more electronic texts appeared in collections such as those of the Electronic Text Center (ETC) at the University of Virginia, in the Oxford Text Archive (OTA) of Humanities Computing at Oxford University, and at the Center for Electronic Text in the Humanities (CETH) at Princeton and Rutgers Universities. It became clear that the humanities computing community would benefit from the experience of catalogers and from cataloging rules that had been practiced in libraries for the purpose of sharing records.

Perhaps in a less obvious manner, catalogers may also apply experience gained from their exposure to the demands of electronic texts in the humanities: the efforts of cataloging and encoding meta-information in electronic texts for the humanities provides direction for cataloging all electronic materials, particularly resources on the Internet. According to Marko (1994), Head of the Monograph Cataloging Division at the University of Michigan Library, catalogers may face a situation similar to that of the ice industry at the advent of refrigeration unless cataloging practices adapt to the changing environment brought about by electronic formats. The Library of Congress (1996a) recommends that catalogers prepare for the future of organizing for access to digital libraries by finding

> ways to expand the use of metadata that forms part of the digital object . . ., include it on digital resources and develop mechanisms for integrating different forms of metadata (MARC, TEI, etc.). Although metadata efforts are more advanced for digital text material (such as those employing the TEI header), other digitized resources (such as text bit-mapped images) could also benefit from metadata schemes.

THE TEI HEADER

The TEI header is attached to an electronic document; it is a label containing directions about the document's logical structure. It may also function independently from the electronic text it pertains to, as a vehicle for meta-information about the electronic material. It is this capacity of producing independent records of meta-information following prescribed rules that make the TEI header of interest to catalogers: in effect, the rules designed by the TEI serve as guidelines for using SGML for the purpose of describing electronic documents in a manner which is pertinent both to the requirements of electronic forms and to those of more traditional cataloging.

The TEI header includes four parts, only one of which is mandatory, but all pertain to issues specific to cataloging electronic forms:

- file description and sources from which the text was derived (mandatory)
  print, electronic
- encoding system which has been applied (optional)
- profile description (optional)
  class, keywords, subject headings
- revision history (optional). (Burnard 1995, part 4.2)

Although only the file description is mandatory, the other parts are highly recommended, especially in the case of independent TEI headers because they contain the possibility of including information that is difficult to describe using AACR2 rules for computer files (Horowitz & Palowitch, 1996). Appendix A presents an example of a TEI header written for the electronic version of Martin Dillon's *Assessing Information on the Internet: Toward Providing Library Services for Computer-Mediated Communication* (Vizine-Goetz, 1995b). Appendix B presents the same record encoded with USMARC (Vizine-Goetz, 1995b).

ANALYSIS OF A TEI HEADER

In the example of a TEI header proposed in Appendix A, the following elements as mentioned above can be seen: the file description (<fileDesc>), the encoding description (<encodingDesc>) shown as non-applicable, the profile description (<profileDesc>), and the revision history (<revisionDesc>) also shown as non-applicable.

General Remarks

The file-description element and the profile-description element are the most familiar to a traditional cataloger since they contain the metadata that forms the basis of a catalog record. The file-description contains descriptive information such as title statement, publication information, and other information. It also contains the source-description element which allows encoding information about the physical text or sources from which the electronic text has been derived. The profile description also contains subject headings and classification schemes according to which the electronic text may be assigned a call number or accession number. The profile-description element may also contain information necessary to humanistic studies ranging from the human languages in which the text has been written to information about the social context.

Subject Headings and Classification

This example of a TEI-encoded document contains Library of Congress Subject Headings (LCSH) and classification codes according to the Dewey Decimal Classification and the Library of Congress Classification schemes. Not all TEI-encoded documents do so (see Appendix C) because the profile-description element is not mandatory, but only highly recommended according to the TEI scheme. Appendix A declares in the <profileDesc> tag that it uses the keyword scheme LCSH, and attributes the following keywords: Internet (Computer Network), Cataloging of computer files, Information networks, Libraries and the sub-heading Communication systems, Information Storage and Retrieval systems, and Library information networks.

The presence or absence of subject headings within a TEI-encoded document should not, however, constitute a criteria for judging the usefulness of the TEI scheme. The profile description allows for the possibility of including subject headings, and the inclusion depends on the purposes for which the electronic document has been created. For instance, the electronic texts found at the ETC at the University of Virginia use their own sets of keywords for describing the content of an electronic text (Gaynor, 1994). In place of the LCSH keyword scheme in the previous examples, Gaynor proposes using a proprietary ETC scheme:

```
<profDesc>...
    <txtClass>
        <keywords scheme=ETC>non-fiction; essays
        </keywords>
    </txtClass>
</profDesc>
```

In Appendix C, the profile description for Zora Neale Hurston's *Their Eyes Are Watching God*, available

from the Center for Electronic Texts in the Humanities (CETH), only specifies languages and does not mention text class and classification scheme. Although the use of LCSH is recommended in certain instances, it may not always be appropriate. A TEI header allows for the possibility of declaring one's own scheme of subject headings or using standardized schemes such as LCSH.

Appendix A declares the following classifications in the <profileDesc> element: Dewey Decimal Classification 004.67 and Library of Congress Classification TK5105.875.I57. Examples from CETH and ETC provide no space for such a classification scheme (Gaynor, 1994).

File Description

The File Description element of the TEI-header includes the following fields:

- title statement
- edition statement
- extent
- publication statement
- series statement
- notes statement
- source description

Each of these fields may be translated to the fields of a USMARC record, and OCLC has announced a prototype program (Spectrum) which allows automatic translation (Vizine-Goetz, 1995a). In the example in Appendix A, the fields <extent> and <series statement> are empty because they are unknown or not applicable to the particular electronic text to be cataloged. Gaynor's example of a TEI-header for ETC includes the size of the electronic file in kilobytes in <extent>. The example in Appendix A includes the size of the electronic file in <Note Statement>.

The source description describes the physical item and various sources from which the computer file is derived. It is often the case that electronic texts such as those available from CETH and ETC are digitized from one or several printed books or manuscripts, and it is very important for the research use of these texts that the user knows exactly which edition he or she is studying. In addition, some electronic archives and repositories of ancient texts share their records: the source description of some texts from ETC indicate that the source is another computer file from the electronic repository at the Oxford Text Archive.

The source description often gives the full bibliographic record of the printed sources, and duplicates the

format of the first six fields of the file description. Regardless of how much the first six fields of the file description resemble the source description fields in content, the user of an electronic text must not forget that the source description contains meta-information about the physical sources, whereas the file description contains meta-information for the electronic record itself.

In the case of electronic documents that may never have appeared in print, such as World Wide Web pages, the source description field need only contain the indication "original" and nothing else (c.f., Appendix A). The source description field may concern itself with the requirements of intellectual property for digitization of print sources.

It should also be noted that the information that may be translated into fields 245 (Title Statement), 260 Publication, Distribution, etc.), and 650 (Subject-added entry) of a MARC record for an electronic document must be taken from the first six fields of the file description, as well as from the profile description, but not from the source description, as this would duplicate the record for the printed text and not generate one for the electronic text.

Authority Control

Many catalogers have pointed out that early electronic texts and those created without regard for cataloging rules do not practice authority control. The rules for encoding information into a TEI header do not prescribe the use of AACR2 for writing information into the fields. Therefore the title and author fields in the <title statement> element of file description may be spelled and capitalized according to the encoder's fancy or to the policies in effect in his or her organization. This may render access to information from these fields difficult.

Fortunately, efforts to create electronic texts on a large scale, such as those at CETH and ETC, have been working in partnership with catalogers. Gaynor (1994) describes how the ETC developed a set of local guidelines for entering information concerning the author, publication, and edition statements in the TEI headers used at ETC, with the purpose of making the TEI header as congruent to a MARC record as possible. A particular effort was made for the completeness and accuracy of publication information (USMARC field 260), often a volatile area for electronic information. Gaynor reports that the indexing and retrieval software provided flexibility for the author-related fields, with the results that TEI headers at ETC do not have to conform to AACR2. Thomas Jefferson as an author's name may be entered

in the TEI header and searched as: Jefferson; Thomas Jefferson; Jefferson, Thomas; Thos. Jefferson, and T. Jefferson; according to how it is spelled in the printed source.

Even a TEI-header created by OCLC, such as the one in Appendix A, does not format the author field in the title statement according to AACR2. The author's names are listed as "Martin Dillon, Erik Jul" and not "Dillon, Martin; Jul, Erik".

CONGRUENCE OF THE TEI HEADER WITH USMARC AND FUTURE DIRECTIONS

The attempts made by OCLC and at the ETC for translating a TEI header into a MARC record reveal that a number of issues need further discussion and elaboration before satisfactory answers may be found (Gaynor, 1994; Vizine-Goetz, 1995a). One is that the granularity of a TEI header must be refined in order to allow effective flow of data to a MARC record. Most TEI headers only specify a general author field, which may contain several authors, and distinguish between author as person, corporate author, and conference proceedings only occasionally. This makes the automatic assignment of MARC tags difficult in the case of a corporate body as main entry for texts encoded with TEI. This lack of granularity does not render TEI invalid, because the TEI guidelines are flexible enough to include such refinements. But it will depend on each individual creator and distributor of electronic texts, whether organization, library, or educational institution, to ensure that the necessary granularity is present.

It is unclear whether there is space in the current design of MARC fields and subfields for the revision history of the electronic file, a part of the TEI header. MARC field 856 (Electronic Location and Access) allows for a wide range of information regarding access to an electronic file, but does not appear to include an indicator for revision history (Library of Congress, 1997). More investigation is necessary to ascertain if the leader field 005 (Date and Time of Last Transaction) conveys similar kind of information as the revision history field in a TEI header (Library of Congress, 1996b).

A second issue that needs clarification is what constitutes an original version, an edition statement, and a publication status for an electronic document. To what extent does the digitization process for the creation of an electronic document constitute a new intellectual work, a new edition of an existing work, or simply a new format of the same work? For example, TEI encoders at ETC have felt strongly enough about the novelty and value of their work to consider that—at least for texts that did not come from commercial providers—an electronic text is a new intellectual work, and they have registered the University of Virginia Library as publisher (MARC 260). Even more specific, Appendix A proposes the Office of Research at OCLC as publisher.

Questions concerning the publication status of a document which arise with the cataloging of electronic documents call for an agreement between catalogers and TEI encoders because they affect how information is represented in the publication statement, series statement, and notes statement both of the TEI header and of a MARC record. This problem is not one of congruence between MARC and TEI, but a common problem that appears in both schemes and therefore must be resolved in accord.

Projects now underway at the ETC and CETH have shown that the use of a TEI header for the cataloging of electronic documents in general and electronic texts in particular is a fruitful endeavor. TEI introduces minimum standards which may be refined and adapted to suit the specific requirements of electronic-texts creation and distribution. Both partners in the design of TEI headers (catalogers and humanities computing encoders) have found that a TEI header may accommodate descriptive information which is difficult to encode into a MARC record, such as the source description and the revision history of an electronic document. They have also found that a sound measure of authority control is necessary for the file description field of the TEI header although TEI guidelines do not prescribe it. The TEI header and a MARC record differ significantly because they were created for different purposes, but in actual usage, the convergence that exists between the two may be exploited with numerous advantages. When adapted to allow better congruence, TEI headers and the MARC structure offer possibilities for cataloging all sorts of electronic documents and not only electronic textual material.

APPENDIX A

TEI header for Martin Dillon's *Assessing information on the Internet:*
Toward providing library services for computer-mediated communication. (Vizine-Goetz, 1995b)

\<teiHeader>
\<fileDesc>
  \<titleStmt>
    \<title>Assessing Information on the Internet: Toward Providing Library Services for Computer Mediated Communication\</title>
    \<author>Martin Dillon\</author>
    \<author>Erik Jul\</author>
    \<author>Mark Burge\</author>
    \<author>Carol Hickey\</author>
  \</titleStmt>
  *\<editionStmt>*NA\</editionStmt>
  *\<extent>*NA\</extent>
  \<publicationStmt>
    \<publisher>OCLC Online Computer Library Center, Inc., Office of Research\</publisher>
    \<address>6565 Frantz Road Dublin, Ohio 43017-3395\</address>
    \<date>1994\</date>
  \</publicationStmt>
  *\<seriesStmt>*NA\</seriesStmt>
  \<notesStmt>
    \<note> 856 7 $u http://www.oclc.org/oclc/menu /reschdoc.htm $z For an introductory page to an electronic version of: Assessing information on the Internet $2 http\</note>
    \<note> 856 1 $a ftp.rsch.oclc.org $d ftp/pub/internet_resources_project /report
$f cover.ps $s 9679 bytes
$f internet.ps $s 257990 bytes
$f appenda.ps $s 84957 bytes
$f appendb.ps $s 66017 bytes
$f appendc.ps $s 37973 bytes
$f appendd.ps $s 46106 bytes
$f appende.ps $s 351941
$u ftp://ftp.rsch.oclc.org/pub /internet_resources_project/report
$z These files are in PostScript format. You may read them online if you have a PostScript viewer. Otherwise, load them to disk and print them on a PostScript printer\</note>
    \<note> 856 1 $a ftp.rsch.oclc.org $c Must be decompressed with Unix uncompress $c Must be untarred with Unix tar -xvf $d ftp/pub/internet_resources _project/report
$f report.ps.tar.Z $s 312328 bytes
$u ftp://ftp.rsch.oclc.org/pub /internet_resources_project/report /report.ps.tar.Z\</note>
  \</notesStmt>

\<sourceDesc>
  \<biblFull>
    \<titleStmt>
      \<author>;Martin Dillon ... [et al.] \</author>
      \<title>Assessing information on the Internet : toward providing library services for computer-mediated communication\</title>
    \</titleStmt>
    \<editionStmt>
      \<edition>;NA\</edition>
    \</editionStmt>
    \<extent>1 v. (various pagings) : ill. ; 29 cm.\</extent>
    \<publicationStmt>
      \<resp>\<role>publisher\</role> \<name>OCLC\</name>\</resp>
      \<place>Dublin, Ohio\</place>
      \<idno type='OCLC'>27635027\</idno>
      \<date>1993\</date>
    \</publicationStmt>
    \<sourceDesc>No source: this is an original work\</sourceDesc>
  \</biblFull>
  \</sourceDesc>
\</fileDesc>

\<encodingDesc>NA\</encodingDesc>

\<profileDesc>
  \<textClass>
    \<keywords scheme=LCSH>
      \<list>
      \<item>Internet (Computer network)\</item>
      \<item>Cataloging of computer files\</item>
      \<item>Information networks\</item>
      \<item>Computer networks\</item>
      \<item>Libraries---Communication systems \</item>
      \<item>Information storage and retrieval systems\</item>
      \<item>Library information networks\</item>
      \</list>
    \</keywords>
    \<classCode scheme=DDC20>004.67\</classCode>
    \<classCode scheme=LCC>TK5105.875.I57
    \</classCode>
  \</textClass>
\</profileDesc>

\<revisionDesc>NA\</revisionDesc>
\</teiHeader>

APPENDIX B

MARC-format record translated from TEI header
in Appendix A. (Vizine-Goetz, 1995b)

090	RK5105.875.I57
092	00467 $2 20
100	Martin Dillon
245	Assessing information on the internet $h [computer file] : toward providing library services for computer mediated communication
260	Dublin, Ohio : $b OCLC Online Computer Library Center, Inc., Office of Research $c 1994
650	Internet (Computer network)
650	Cataloging of computer files
650	Information networks
650	Libraries---Communication systems
650	Information storage and retrieval systems
650	Library information networks
700	Erik Jul
700	Mark Burge
700	Carol Hickey
856 7	$u http://www.oclc.org/oclc/menu.reschdoc.htm
856 1	$a ftp.rsch.oclc.org $d ftp/pub/internet_resources_project/report $f cover.ps $s 9679 bytes $f internet.ps $s 257990 bytes $f appenda.ps $s 84957 bytes $f appendb.ps $s 66017 bytes $f appendc.ps $s 37973 bytes $f appendd.ps $s 46106 bytes $f appende.ps $s 351941 bytes
856 1	$a ftp.rsch.oclc.org $s ftp/pub/internet_resources_project/report $f report.ps.tar.Z $s 312328

APPENDIX C

Profile description for Zora Neale Hurton's
Their Eyes Are Watching God. (CETH, 1997)

<TEI2>
<TEI HEADER>
<FILE DESCRIPTION>
   <TITLE STATEMENT>
      <TITLE>Their Eyes Were Watching God</TITLE>
      <AUTHOR>Zora Neale Hurston</AUHTOR>
      <EDITOR>Anthony Lioi</EDITOR>
   </TITLE STATEMENT>

   <PUBLICATION STATEMENT>
      <DISTRIBUTOR><NAME>Center for Electronic Texts in the Humanities</NAME>
      <ADDRESS>169 College Avenue, New Brunswick, NJ 08903</ADDRESS>
      </DISTRIBUTOR>
      <AVAILABILITY> Freely available for non-commercial us when distributed with this header intact.</AVAILIBILITY>
   </PUBLICATION STATEMENT>

<SOURCE DESCRIPTION>
   <SOURCE><BIBL><AUTHOR>Zora Neale Hurston</AUTHOR>
      <TITLE>Their Eyes Were Watching God</TITLE>
      <EDITION>Perennial Library Edition</EDITION>
      <PUBLISHER>Harper & Row</PUBLISHER>
      <PUBPLACE>New York</PUBPLACE>
      <DATE>1990</DATE></BIBL>
   </SOURCE DESCRIPTION>

<ENCODINGDESC><PROJECTDESC>This text was prepared as a TEI pilot project.
   </PROJECTDESC>
   <TAGSDESC>TEI tags declaration
      <RENDITION>Words and phrases surrounded by the TEI tag "seg" should be rendered in a color which distinguishes them from words and phrases surrounded by the TEI tags "ref" and "sic."</RENDITION>
      <USAGE> The TEI tag "seg" appears when Hurston employs a trope
      of narration-as-kissing.</USAGE>
   </TAGSDESC>
</ENCODINGDESC>

<PROFILEDESC><LANGUAGES>Standard Written American English
      Black English Vernacular of Florida in the 1920's
      A hybrid idiolect of Standard and BEV
   </LANGUAGES>
</PROFILEDESC>
</TEI HEADER>

<TEXT>
<BODY>

<HEAD>Chapter 1 </HEAD>
 <P>Ships at a distance have every man's wish on board. <NOTE>Hurston's opening refers to the scene in Chapter X of Frederick Douglass' Narrative in which the narrator gazes out at the ships passing through Baltimore Harbor and thinks of freedom: "Our house stood within a few rods of Chesapeake Bay, whose broad bosom was ever white with sails from every quarter of the habitable globe. Those beautiful vessels, robed in purest white, so delightful to the eye of freedmen, were to me so many shrouded ghosts, to terrify and torment me with thoughts of my wretched condition. I have often, in the deep stillness of a summer's Sabbath, stood all alone upon the lofty banks of that noble bay, and traced, with saddened heart and tearful eye, the countless number of sails moving off to the mighty ocean. The sight of these always affected me powerfully." Frederick Douglass Narrative of the Life of Frederick Douglass, an American Slave. Written by Himself in The Classic Slave Narratives edited and introduced by Henry Louis Gates, Jr New York New American Library 1987 p. 293.</NOTE>

For some they come in with the tide. For others they sail forever on the horizon, never out of sight, never landing until the Watcher turns his eyes away in resignation, his dreams mocked to death by Time. That is the life of men.</P>

<P>Now, women forget all those things they don't want to remember, and remember everything they don't want to forget. The dream is the truth. Then they act and do things accordingly.</P>

REFERENCES

Burnard, L. (1995, July). *Text encoding for information interchange: An introduction to the Text Encoding Initiative* (TEI Document No. TEI J31). Available from http://www.uic.edu/orgs/tei/info/teij31

Burnard, L., & Sperberg-McQueen, C. M. (Eds.). (1994, April 8). *Guidelines for electronic text encoding and interchange. TEI P3. Chicago: Text Encoding Initiative*. Available from http://www.uic.edu/orgs/tei/p3/doc/p3.html

Center for Electronic Texts in the Humanities (CETH). Rutgers and Princeton Universities. (1997). Zora Neale Hurston's *Their Eyes Are Watching God*'s Chapter 1: An Annotated Electronic Version of the Harper Perennial Edition. Anthony Lioi (Ed.). Available from http://www.ceth.rutgers.edu/IntroductoryMaterial/SGMLsamples/Hurston/hurston.sgml

Electronic Text Center (ETC). University of Virginia. (1998). Available from http://etext.lib.virginia.edu

Gaynor, E. (1994). Cataloging electronic texts: The University of Virginia Library experience. *Library Resources and Technical Services*, *38*(4), 403-412.

Horowitz, L., & Palowitch, C. (1996). Meta-information structures for networked information resources. *Cataloging and Classification Quarterly*, *21*(3/4), 109-130.

Library of Congress. (1996a). *Organizing the Global Digital Library*: Proceedings of the First Organizing the Digital Library Conference held in Washington, D.C. 11 December, 1995. Available from gopher://marvel.loc.gov/00/loc/conf.meet/gdl

Library of Congress. (1996b). *The USMARC formats: Background and principles*. Available from http://lcweb.loc.gov/marc/96principl.html

Library of Congress. (1997, August). *Guidelines for the Use of Field 856*. Available from http://lcweb.loc.gov/marc/856guide.html

Marko, F. L. (1994). *Technology shift and the impact on cataloging: A view from the University of Michigan*. Available from http://lcweb.loc.gov/catdir/semdigdocs/marko.html

Vizine-Goetz, D. (1995a). Office of Research project develops tools for describing and accessing Internet resources. *OCLC Newsletter, 213* (January-February): 13-16.

Vizine-Goetz, D. (1995b). *Cataloging productivity tools*. Available from http://www.oclc.org/oclc/research/publications/review94/part1/1spectr.html

ADDITIONAL READINGS

Burnard, L., & Sperberg-McQueen, C. M. (1995). *TEI Lite: An introduction to Text Encoding for Interchange*. TEI U5. Available from http://www-tei.uic.edu/orgs/tei/intros/teiu5.html

Gorman, M., & Winkler, P. W. (Eds.). (1988). *Anglo-American Cataloguing Rules* [AACR2] (2nd ed.). Chicago: American Library Association.

Library of Congress. (1996). *Organizing the Global Digital Library II (OGDL II) and Naming Conventions*: Proceedings of the Second Organizing the Digital Library Conference held in Washington, D.C. 21-22 May 1996. Available from http://lcweb.loc.gov/catdir/ogdl2/ogdlhome.html