



## **Committee on Cataloging: Description & Access**

# **Task Force on Metadata and the Cataloging Rules**

## **Final Report (continued)**

### **Table of Contents**

#### ***Executive Summary***

Introduction

Metadata and Cataloging

The TEI Header and the Cataloging Rules

Dublin Core Metadata and the Cataloging Rules

Encoded Archival Description: Summary Report

Conclusions

Recommendations

Bibliography

Appendix: Cataloging Problems with Web Sites

## **The TEI Header and the Cataloging Rules**

**By Jackie Shieh, *University of Virginia Library***

[The description in this report is based on the *Guidelines for Electronic Text Encoding and Interchange (TEI P3)* / edited by C.M. Sperberg-McQueen and Lou Burnard. — Chicago ; Oxford : Text Encoding Initiative, c1994. Also available [ASCII and SGML

versions] <http://www.uic.edu/orgs/tei/p3/>; [www page] <http://etext.virginia.edu/TEI.html>

This report is based on extensive experience with TEI headers at the University of Virginia. It thus represents a typical, but not necessarily comprehensive, presentation of issues.]

## **Background**

The object of this section is to evaluate TEI header metadata as a source of cataloging data for records based on the *Anglo-American cataloging rules*, using the University of Virginia's Electronic Center project as the test case. The first parts will describe the TEI header data elements and structure. Subsequent parts will analyze the data elements within the context of AACR2r and its transfer syntax, USMARC Format for Bibliographic Data. A copy of the UVA TEI/AACR/MARC mapping is included. Because there is no Network Development and MARC Standards Office crosswalk for the TEI header/MARC, analysis of mapping between these is based on the UVA model.

### *What is a TEI header?*

The Text Encoding Initiative (TEI) was established in 1987 under the joint sponsorship of the Association for Computers and the Humanities, the Association for Computational Linguistics, and the Association for Literary and Linguistic Computing. The impetus for the project originated from the humanities computing community seeking a common encoding scheme for complex textural structures to reduce the diversity of the existing encoding practices. The focus of the use for TEI was later broadened to meet the varied encoding requirements of any discipline or application.

The TEI Guidelines do not represent a static finished work, but rather one which will evolve over time with the active involvement of its community of users to ensure that the TEI Guidelines become and remain useful in all sorts of work with machine-readable texts.

Every TEI-conformant text contains an "electronic title page" – the TEI header. The header is not the substitute of a printed title page, rather it is an encoder's interpretation of the bibliographic

information. The information contained in a TEI header describes and documents an encoded work for the text itself, its source, its encoding process and revisions. The documentation is vital not only for the scholars using the materials, but also the catalogers in libraries and archives. The description of the text and its encoding provide an electronic analog to the title page of a printed work, or an equivalent to the codebook or introductory manual accompanying the electronic data sets.

### *What is in a TEI header?*

A TEI header can be a large and complex or a simple object. Every header must carry a set of tagged descriptive elements, <teiHeader> which contains four major components (only the <fileDesc> element is required in all TEI headers; the others are optional):

<teiHeader>

<fileDesc> ... </fileDesc>

- containing a full bibliographic description of an electronic file.

<encodingDesc> ... </encodingDesc>

- documenting the relationship between an electronic text and the source or sources from which it was derived.

<profileDesc> ... </profileDesc>

- providing a detailed description of non-bibliographic aspects of a text, specifically the languages and sublanguages used, the situation in which it was produced, the participants and their setting.

<revisionDesc> ... </revisionDesc>

- summarizing the revision history for a file.

</teiHeader>

*What is contained in the four components of a TEI header and how can these be used as sources of information for cataloging?*

[The following description for the four components is based on the teilight.dtd (TEI P3) used at the University of Virginia (UVa), Electronic Text Center. The description conforms to the document found in UVa's ETEXT Center site <http://etext.virginia.edu/TEI.html>. For complete description on TEI encoding initiatives, please refer to TEI home site, <http://www.uic.edu:80/orgs/tei/>.]

## **The File Description — <fileDesc>**

This segment closely resembles the bibliographic description in structure for the electronic text. The elements are modeled after the existing standards in library cataloging. Therefore, this segment should provide enough information allowing users to give standard bibliographic references to the electronic text, and catalogers to catalog it. This is a mandatory field (a minimal <fileDesc> must contain at least the following structure: <titleStmt>, <publicationStmt>, <sourceDesc>).

### **<fileDesc>**

**\*<titleStmt> ... </titleStmt>**

AACR2 9.1B-F; 9.7B4; 21-22, 24.  
USMARC 100; 110; 245 \$a,b,n,p,h.  
(information about the title of a work and those responsible for its intellectual content)

**<editionStmt> ... </editionStmt>**

AACR2 9.2B; USMARC 250.  
(information relating to one edition of a text)

**<extent> ... </extent>**

AACR2 9.3B; USMARC 256.  
(information on the approximate size of the electronic text as stored on some carrier medium, specified in any convenient units)

**\*<publicationStmt> ... </publicationStmt>**

AACR2 9.4C-D; USMARC 260 \$a,b,c.  
(information concerning the publication or distribution of an electronic text)

**<seriesStmt> ... </seriesStmt>**

AACR2 9.6B; USMARC 440.  
(information about the series, if any, to which the electronic publication belongs)

**<notesStmt> ... </notesStmt>**

AACR2 9.7; USMARC 5XX.

(information about a text additional to that recorded in other parts of the bibliographic description)

**\*<sourceDesc> ... </sourceDesc>**

AACR2 9.7B7; USMARC 500, 534.

(information on a bibliographic description of the copy text(s) from which an electronic text was derived or generated)

**</fileDesc>**

## **The Encoding Description — <encodingDesc>**

This segment specifies the methods and editorial principles which govern the transcription or encoding of the text in hand and may also include sets of coded definitions used by other components of the header. This field is not required, but highly recommended.

**<encodingDesc>**

**<projectDesc> ... </projectDesc>**

AACR2 9.7B6; USMARC 500.

(Details the purpose for which the encoded electronic file, together with any other relevant information concerning the process by which it was assembled or collected)

**<editorialDesc> .... </editorialDesc>**

AACR2 9.7B8; USMARC 516.

(Details of editorial principles and practices applied during the encoding of the electronic text)

**<refsDecl> ... </refsDecl>**

AACR2 9.7B8; USMARC 500, 516.

(Details the construction of the canonical references)

**<classDecl> ... </classDecl>**

N/A in AACR2 or USMARC.

(Containing one or more taxonomies defining any classificatory codes used elsewhere in the text, such as Library of Congress Subject Headings)

**</encodingDesc>**

## **The Profile Description — <profileDesc>**

This segment provides a detailed description of non-bibliographic aspects of a text, specifically the languages and sublanguages

used, the situation in which it was produced, the participants and their setting. This field is optional.

**<profileDesc>**

**<creation> ... </creation>**

(information about the creation of the electronic text)

**<langUsage> ... </langUsage>**

AACR2 9.7B2; USMARC 041, 546.

(information on the languages, sublanguages, registers, dialects etc. represented within the electronic text)

**<textClass> ... </textClass>**

N/A in AACR2; USMARC 6XX

(information describing the nature or topic of the electronic text in terms of a standard classification scheme, such as LCSH, thesaurus, etc.)

**</profileDesc>**

## **The Revision Description — <revisionDesc>**

This segment provides a detailed log on each change made to an electronic text. It is recommended to give changes in reverse chronological order, most recent first. This field is optional.

**<revisionDesc>**

**<change>**

**<date> ... </date>**

AACR2 9.7B; USMARC 9XX.

(containing a date in any format)

**<respStmt> ... </respStmt>**

(statement of responsibility for someone responsible for the intellectual content of the text, edition, recording, or series, where the specialized elements for authors, editors, etc. do not suffice or do not apply)

**<item> ... </item>**

(indicates what change was made; it can range from a simple phrase to a series of paragraphs)

**</change>**

**</revisionDesc>**

## TEI Header Metadata's Support of the Four User Tasks

The TEI header metadata supports the four user tasks as described in the document *Functional Requirements for Bibliographic Records* in varying degrees. The TEI header tag (<teiHeader>), contains four major components. Only the File Description (<fileDesc>) is mandatory. Within this element, the Title Statement (<titleStmt>), the Publication Statement (<publicationStmt>) and the Source Description (<sourceDesc>) are the required constituents. The amount of encoding in a header depends on both the nature and intention of the text. Thus, the header can be a very simple or complex object.

Therefore, the satisfaction on the support level for the user tasks may range from one extreme to another. The following paragraphs denote the practice of TEI header at the University of Virginia (UVA). UVA's header is intended to optimize the four user tasks by employing the established cataloging principles. A TEI header is edited by the cataloging staff to provide bibliographic information comparable to a full-level cataloging record.

- to **find** entities that correspond to the user's stated search criteria (i.e., to locate an entity in a file or database as the result of a search using an attribute or relationship of the entity);

A TEI header contains elements which help user distinguish search activity for relevancy on primary and secondary searches. The <title> and <author> elements in <fileDesc> are most likely the primary search feature in seeking for the electronic equivalent of a printed source. The content standard employed in the <author> tag utilizes the national authority data file which ensures the quality of the <author> tag and results in a higher and correct retrieval rate by the user.

The controlled vocabularies used in the <profileDesc> <textClass> <keyword scheme="LCSH"> element, serving as secondary search feature, can be used to further qualify a desired search output.

- to **identify** an entity (i.e., to confirm that the entity described corresponds to the entity sought, or to distinguish between two or more entities with similar characteristics);

The <fileDesc> <sourceDesc> element provides extensive information for the printed source from which the electronic version was created. This segment includes

not only the original title for the monographic or serial entry (<title level="a,m,j,u">), but also edition <editionStmnt>, series <seriesStmnt>, and other relevant note statements <notesStmnt> which enable the user to identify the resources retrieved and distinguish similar items which may not be as desirable for his research purposes.

- to **select** an entity that is appropriate to the user's needs (i.e., to choose an entity that meets the user's requirements with respect to content, physical format, etc., or to reject an entity as being inappropriate to the user's needs);

The extensive data description relating to the electronic version in the <fileDesc> <notesStmnt> element efficiently provides the necessary information for the user to determine the relevancy of the retrieved item.

- to **acquire** or obtain access to the entity described (i.e., to acquire an entity through purchase, loan, etc., or to access an entity electronically through an on-line connection to a remote computer)

The <idno> provides unique identifier for the sought resource. However, it is not used as the primary retrieving feature for the resource. The <availability> tag contains active URL link to the text which is fully supported in a networked environment.

The quality of a TEI header varies greatly depending on the original encoder. When a header is produced based on the cataloging practices and principles, the user retrieval satisfaction is enhanced. The content standards (ISBD, AACR2, and authority control for names and subject headings) deployed in a header ensures higher level of retrieval accuracy. When the integrity and consistency of a cataloging record is not compromised, the performance of metadata in the support of the Four User Tasks will be as expected and assured.

UVA's TEI header to a greater extent, has intended to contain more than in a cataloging record. The levels of information found in a header are vital, especially the encoding history and document treatment process, etc. Though, the recorded information may not be as relevant as part of a bibliographic data in a cataloging record, nevertheless it remains an intricate part of the data. One metadata does not fully replace the other in terms of its intended and potential usage.



## Cataloging Rules in AACR2 Chapter 9 for Computer Files

**9.0B1. Chief source of information.** The chief source of information for computer files is the title screen(s).

If there is no title screen, take the information from other formally presented internal evidence (e.g., main menus, program statements, first display of information, the header to the file including "Subject" lines, information at the end of the file). In case of variation in fullness of information found in these sources, prefer the source with the most complete information.

If the computer file is unreadable without processing (e.g., compressed file, printer-formatted file), take the information from the file after it has been uncompressed, printed out, or otherwise processed for use.

If the information required is not available from internal sources, take it from the following sources (in this order of preference)

- the physical carrier or its labels
- information issued by the publisher, creator, etc., with the file (sometimes called "documentation")
- information printed on the container issued by the publisher, distributor, etc.

If the item being described consists of two or more separate physical parts, treat a container or its permanently affixed label that is the unifying element as the chief source of information if it furnishes a collective title and the formally presented information in, or the labels on, the parts themselves do not.

If the information required is not available from the chief source or the sources listed above, take it from the following sources (in this order of preference)

- other published descriptions of the file
- other sources

As an "electronic title page," a TEI header could be considered as the chief source of information for the bibliographic record. It is recommended that the TEI header only be considered as such when the title screen is absent. The header is created by the person who creates the electronic text. This person may or may not be aware of the intricacies of the cataloging rules which govern bibliographic data entry. Further, the header describes the electronic version, not the original from which the electronic version was created and thus does not necessarily record the data as it appears on the title page of the original. Only in rare cases will the header conform to the transcription of the text on which the electronic version is based. As it is pointed out in the

**5.7 Note for Library Cataloguers** from the TEI Guidelines (*TEI P3*, p. 137), “the header is not a library catalogue record, and so will not make all of the distinctions essential in standard library work. It is the intention of the developers, however, to ensure that the information required for a catalogue record be retrievable from the TEI file header, and moreover that the mapping from the one to the other be as simple and straightforward as possible.” In keeping with these points, it is recommended that TEI headers not be considered library catalog record substitutes. They are important as “other formally presented internal evidence” however, and can serve as rich resource records for the creation of library catalog records.

Failure to use *AACR2*-prescribed punctuation and capitalization is another reason not to consider headers as library catalog record substitutes. Sections of the rules are extracted here to help demonstrate the problems.

**9.0B2 Prescribed sources of information.** The prescribed source(s) of information for each area of the description of computer files is set out below. Enclose information taken from outside the prescribed source(s) in square brackets.

AREA	PRESCRIBED SOURCES OF INFORMATION
Title and statement of responsibility	Chief source of information, the carrier or its labels, information issued by the publisher, creator, etc., container
Edition	Chief source of information, the carrier or its labels, information issued by the publisher, creator, etc., container
File characteristics	Any source
Publication, distribution, etc.	Chief source of information, the carrier or its labels, information issued by the publisher, creator, etc., container
Physical description	Any source
Series	Chief source of information, the carrier or its labels, information issued by the publisher, creator, etc., container
Note	Any source
Standard number and terms of availability	Any source

## **9.0C Punctuation**

For the punctuation of the description as a whole, see 1.0C.

For the prescribed punctuation of elements, see the following rules.

## **[Appendix A. Capitalization]**

### **A.4 Title and Statement of Responsibility Area**

**A.4A1.** Capitalize the first word of the title proper, an alternative title, or a parallel title (see also A.4B below). Capitalize other words, including the first word of each element of other title information, as instructed in the rules for the language involved (see also A.4D). See A.20 for the capitalization of names of documents.

The materials of architecture  
The 1919/20 Breasted Expedition of  
the Near East  
Les misérables  
IV informe de gobierno  
Eileen Ford#70146;s more beautiful  
you in 21 days  
Journal of polymer science  
Sechs Partiten für Flöte  
The Edinburgh world atlas, or,  
Advanced atlas of modern geography  
Coppélia, ou, La fille aux yeux  
d'émail  
...  
King Henry the Eighth ; and, The  
tempest

### **A.5 Edition Area**

**A.5A**If an edition statement (or a statement relating to a named revision of an edition) begins with a word or an abbreviation of a word, capitalize it. Capitalize other words as instructed in the rules for the language involved.

Household ed.  
Facsim. ed.  
1st standard ed.  
Neue Aufl.  
Rev. et corr.  
Wyd. 2-gie  
World's classics ed., New ed. rev.

## **A.9 Series Area**

**A.9A1.** Capitalize the title proper, parallel titles, other title information and statements of responsibility of a series as instructed in A.4.

Concertino : Werke für Schul- und  
Liebhaber Orchester  
Jeux visuels = Visual games

**A.9B1.** Do not capitalize a term such as v., no. reel, t., that is part of the series numbering unless the rules for a particular language require capitalization (e.g., noun capitalization in German). Capitalize other words and alphabetic devices used as part of a numbering system according to the usage of the item.

Deutscher Planungsatlas ; Bd. 8  
National standard reference data  
series ; NSRDS-NBS 5

## **A.10 Notes Area**

Capitalize the first word in each note or an abbreviation beginning a note. If a note consists of more than one sentence, capitalize the first word of each subsequent sentence. Capitalize the first word of following introductory wording and a colon (see AACR2 1.7A1)

Title from container  
Facsim. reprint. Originally published:  
London : I. Walsh, ca. 1734

It is not expected that the creator of a TEI header will know or even be aware of the AACR2 standards for prescribing sources, or the knowledge of its punctuation rules relating to each element of the header, and so forth.

The punctuation of a header most commonly follows the pattern of a printed title page, where the first letter of each word excluding prepositions and the initial articles is capitalized in each tagged field. For example:

<title>The Three Old Sisters and the Old Beau</title>

There are problems with recording information about editions, statements of responsibility and so on.

***Uncle Tom's Cabin, or, Life Among the Lowly***

**Harriet Beecher Stowe**  
Penguin Books

New York  
1986

**Note:** This text was entered from the Penguin Classics 1986 version of the first-edition and omits the copyrighted introduction. Penguin in turn uses the type-facsimile of the first edition as established by Kenneth S. Lynn, editor of the Belknap Press edition published by Harvard University Press (1962).

When describing an edition used for the creation of the electronic text, the creator of the header generally interprets, rather than transcribes the information found on the title page (t.p.) or t.p. verso. With editing of the above by a UVA cataloger, the edition statement is more concise and clearer. At the University of Virginia, the cataloging staff are viewed as the quality control unit for its TEI headers. All headers receive full level cataloging including Library of Congress subject heading analysis when applicable.

***Uncle Tom's Cabin, or, Life Among the Lowly***

**Harriet Beecher Stowe**

Viking Penguin  
New York  
c1981 (1986 printing)  
(Penguin classics)

**Note:** Published in The Penguin American Library 1981, reprinted 1982, 1983, 1984, 1985 (twice) and reprinted in Penguin Classics 1986.

**Note:** Originally published: Boston : John Jewett & Co. ;  
Cleveland, Ohio : Jewett, Proctt & Worthington, 1852.

There are instances when various editions (printed and/or machine readable versions) are used to create a single electronic text. The text comes from one edition, illustrations, introduction and other editorial remarks from others. In addition, a header can be just a part of a work, e.g. the introduction, illustration only. The current header setup does not show relationships very well. In the <sourceDesc> for "In" analytics, there is no way for relational tags to be grouped together, for example. When one header has multiple titles, each with its own author, or edition statements, little can be done to tell the relationship of one author tag to one particular title tag and other to others and so on.

```
<title>Title (introduction of a work only)
</title>
<title>Title from which the intro (or
  editorial) is used</title>
<author>Author of intro.</author>
<author>Author(1) of the actual work;
  Author (2)</author>
```

**or**  
<author>Author(2) when applicable</author>

<respStmt>  
  <resp>Illustrator or Editor</resp>  
  <name>Name of the illustrator or editor</name>  
</respStmt>

A statement of responsibility can be broken down into various tags in the <respStmt> segment or not recorded.

translation by A.M. Duncan ; introduction and commentary by E.J. Aiton ; with a preface by I. Bernard Cohen.

<respStmt>  
  <resp>Translator</resp>  
  <name>A.M. Duncan</name>  
  <resp>Commentator</resp>  
  <name>E.J. Aiton</name>  
</respStmt>

The statement of responsibility does not appear as the rules of AACR2 prescribe. This may result in conflicting bibliographic references and more investment on the part of the user to identify works.

Authoritative forms of name are generally not used in the header. For example, Mary E. Wilkins has been established in the personal name authority file as **Freeman, Mary Eleanor Wilkins, 1852-1930**. She also had works under the latter entry. It is not uncommon to find dual personal name entries for the same author in headers for the electronic texts.

<author>Wilkins, Mary E.</author>  
**or**  
<author>Mary E. Wilkins</author>

It is also quite common to find that a literary author used pseudonym(s). Elizabeth Gaskell had a pseudonym which was almost unknown to her readers. When an electronic text was created under the name, **Cotton Mather Mills**, there was little evidence linking the work to Mrs. Gaskell. No connection would have been made to verify whether **Cotton Mather Mills** was indeed **Elizabeth Gaskell** without checking authority files or doing research. After several investigations by cataloging staff, one of the Gaskell experts finally remembered that when the author first began her career as a writer, **Cotton Mather Mills** was indeed her *nom de plume*. Confusion can occur when the creator of the header is not aware of the existence of a name authority file or does not understand the principle of a uniform

name heading. This is a common knowledge to catalogers. Thus, variant name entries for bibliographic records are generated for the same author based on the headers when the cataloger is not involved in the process.

## TEI Header Elements and Corresponding MARC Fields

Unlike HTML, the coding in TEI is case-sensitive.

The version of the TEI header (based on *TEI P3*), teilight.dtd that UVA uses, is comprised of four major sections discussed earlier:

```
<teiHeader>
  I. <fileDesc> ... </fileDesc>
  II. <encodingDesc> ... </encodingDesc>
  III. <profileDesc> ... </profileDesc>
  IV. <revisionDesc> ... </revisionDesc>
</teiHeader>
```

- I. **The File Description** – `<fileDesc>` – contains a full bibliographical description of the online version of the text: title, author, creator of the electronic version, size of the file, date of creation, publisher of the electronic version, and information of the print source from which the electronic version was created.

```
<fileDesc>
  <titleStmt>
    <title> ... </title>
    <author> ... </author>
    <respStmt>
      <resp> ... </resp>
      <name> ... </name>
    </respStmt>
  </titleStmt>
  <extent> ... </extent>
  <publicationStmt>
    <publisher> ... </publisher>
    <pubPlace> ... </pubPlace>
    <idno> ... </idno>
    <availability> ... </availability>
    <date> ... </date>
  </publicationStmt>
  <seriesStmt><p> ... </p></seriesStmt>
```

```

<notesStmt><note> ... </note></notesStmt>

<sourceDesc> <biblFull>

  <titleStmt>
    <title> ... </title>
    <title level="j,m,a,u"> ...
    </title>
    <author> ... </author>

    <respStmt>
      <resp> ... </resp>
      <name> ... </name>

    </respStmt>

  </titleStmt>
  <editionStmt><p> ... </p></editionStmt>
  <extent> ... </extent>

  <publicationStmt>
    <publisher> ... </publisher>
    <pubPlace> ... </pubPlace>
    <idno> ... </idno>
    <availability> ... </availability>
    <date> ... </date>

  </publicationStmt>
  <seriesStmt><p> ... </p></seriesStmt>
  <notesStmt><note> ... </note></notesStmt>

</biblFull> </sourceDesc>

</fileDesc>

```

<b>TEI tags</b>	<b>AACR2 (Chapter 9)</b>	<b>MARC tags</b>
<titleStmt> <title>	9.1B-E 9.7B4	245 a, b, n, p, h 246
<titleStmt> <author>	Main Entry See Chapter 21	100 a, b, c, d, q
<resp> <name>	9.1F	500 700, 710, 711
<extent>	9.3B	256
<publicationStmt> <publisher>	9.4B	260 b
<publicationStmt> <pubPlace>	9.4B	260 a
<idno type>		099
<availability> <p>	9.7B9	500



<availability>	Holdings	856 42
<p>URL:		
<date>	9.4B	Fixed field: Date1 260 c
<seriesStmt> <p>	9.6B1	440
<notesStmt> <note>	9.7B9	500
<biblFull> <titleStmt>	Title added	240
<title>	entry	700 X2 t
<title level="j,m,a,u">		730
		773 t
		534 t
<biblFull> <titleStmt	9.1F; 9.7B6	245 c
<author		
<biblFull>	9.2B	534 b
<editionStmt> <p>		
<biblFull> <extent>	9.3B	534 e
<biblFull>	9.4B	534 c
<publicationStmt>		
<publisher>		
<biblFull>	9.4B	534 c
<publicationStmt>		
<pubPlace>		
<biblFull>	9.4B	534 c
<publicationStmt>		
<date>		
<biblFull>	9.6B1	534 f
<seriesStmt> <p>		
<biblFull>	9.7B9	534 n
<notesStmt> <note>		

II. **The Encoding Description** – <encodingDesc> – describes the process of the normalization of the text during transcription, the encoder resolving ambiguities in the source, the application of the levels of encoding or analysis were applied, etc.

<encodingDesc>

<projectDesc>  
 <p> ... </p>  
 </projectDesc>

```

<editorialDecl>
  <p> ... </p>
</editorialDecl>

<refsDecl>
  <p> ... </p>
</refsDecl>

<classDecl>

  <taxonomy id=LCSH>
    <bibl>
      <title>Library of Congress Subject Headings</title>
    </bibl>
  </taxonomy>

</classDecl>
</encodingDesc>

```

TEI tags	AACR2 (Chapter 9)	MARC tags
<projectDesc> <p>	9.7B6	500
<editorialDecl> <p>	9.7B8	516
<refsDecl> <p>	9.7B8	500; 516

III. **The Profile Description** – `<profileDesc>` – describes the non-bibliographic aspects of the text, the languages used in the text, the situation in which it was produced, the participants, and their setting.

The `<date>` field in the `<creation>` section is vital. OpenText uses this to construct its “Centuries” document structures.

```

<profileDesc>

  <creation>
    <date> ... </date>
  </creation>

  <langUsage>
    <language> ... </language>
  </langUsage>

  <textClass>

    <keywords>
      <term>non-fiction; prose</term>
    </keywords>

    <keywords scheme="LCSH">

```

```

        <term type="Field600">
        Crane, Stephen,$d1871-1900
        $xCriticism and interpretation.
        </term>
    </keywords>
</textClass>
</profileDesc>

```

TEI tags	AACR2 (Chapter 9)	MARC tags
<langUsage> <language>	9.7B2	041; 546
<keywords scheme="LCSH"> <term type="Field6XX">		6XX

IV. **The Revision History Description** – **<revisionDesc>** – provides a history of changes made during the development of the electronic text.

```

<revisionDesc>
    <change>
        <date> ... </date>
        <respStmt>
            <resp> ... </resp>
            <name> ... </name>
        </respStmt>
        <item> ... </item>
    </change>
</revisionDesc>

```

TEI tags	AACR2 (Chapter 9)	MARC tags
<resp> <name> <item>	9.7B	9XX local processing note

## Recommendations

The following section is a list of problems and recommendations on mapping for the TEI and Cataloging community to consider.

### Description of Problems

### Recommendations

## **Author**

Cannot easily map titles and other words associated with a name such as relator term or numeration to appropriate 100 subfields

title: **King of England, Saint**  
qualifier: H. G. (**Herbert George**)  
numeration: Louis, **XVI**

Adding the MARC subfields in the author tag:

```
<author>Charles,$cPrince of  
Wales,$d1948-</author>
```

## **Uniform Title**

There is no data element for uniform title which means that all entries for a single work may not be together.

Create a separate title tag to allow mapping, otherwise manual input is required.

*Adventures of Tom Sawyer. 1978*  
*The complete adventures of Tom Sawyer and Huckleberry Finn*

## **Title**

The UVA mapping for alternative titles (246) is based on the presence of additional title information preceded by “or”, a colon, or a semicolon.

It may not be necessary to create a 246(s) from a subtitle(s), or additional title tracing for 740(s) manual input is needed here.

*Uncle Tom's cabin, or, Life among the lowly*  
*A tent in agony : a Sullivan County sketch*

## **Statement of Responsibility**

There is no data element in

Make “Byline” tag which is

the header for statements of responsibility as they appear in the chief source of information (245 \$c)

available in the text structure valid for header.

as told by John Seelye  
translation by A.M.  
Duncan ; introduction and  
commentary by E.J. Aiton ;  
with a preface by I. Bernard  
Cohen.

In these examples, “as told  
by”, “translation by”,  
“introduction and  
commentary by” and “with  
preface by” will not be  
included in the header.

## **Next Section**