

TEI Metadata and Manuscript Description

James Cummings
@jamescummings

12 November 2014

The <teiHeader>

The TEI header was designed with two goals in mind

- needs of bibliographers and librarians trying to document 'electronic books'
- needs of text analysts trying to document 'coding practices' within digital resources

The result is that discussion of the header tends to be pulled in two directions...

The Librarian's Header

- Conforms to standard bibliographic models, using similar terminology
- Organized as a single source of information for bibliographic description of a digital resource, with established mappings to other such records (e.g. MARC, EAD, etc. or embedding in containers like METS)
- General consensus on 'Best Practice for TEI in Libraries' is available from TEI-LIB SIG
- Pressure for greater and more exact constraints to improve precision of description: preference for structured data over loose prose

Everyman's Header

- Gives a polite nod to common bibliographic practice, but has a far wider scope
- Supports a (potentially) huge range of very miscellaneous information, organized in fairly ad hoc or individualistic ways
- Many different codes of practice in different user communities
- Unpredictable combinations of narrowly encoded documentation systems and loose prose descriptions

TEI Header Structure

The TEI header has four main components:

- **<fileDesc>** (file description) contains a full bibliographic description of an electronic file.
- **<encodingDesc>** (encoding description) documents the relationship between an electronic text and the source or sources from which it was derived.
- **<profileDesc>** (text-profile description) 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. (just about everything not covered in the other header elements)
- **<revisionDesc>** (revision description) summarizes the revision history for a file.

Only **<fileDesc>** is required; the others are optional.

Example Header: Minimal required header

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>A title?</title>
    </titleStmt>
    <publicationStmt>
      <p>Who published?</p>
    </publicationStmt>
    <sourceDesc>
      <p>Where from?</p>
    </sourceDesc>
  </fileDesc>
</teiHeader>
```

More Usual Header Structure

```
<teiHeader>
  <fileDesc>
    <titleStmt>
<!-- title/author/editor information -->
    </titleStmt>
    <publicationStmt>
      <p>
<!-- Publication Information -->
      </p>
    </publicationStmt>
    <sourceDesc>
      <p>
<!-- Metadata about the source -->
      </p>
    </sourceDesc>
  </fileDesc>
  <encodingDesc>
    <p>
<!-- Description of the encoding -->
    </p>
  </encodingDesc>
  <profileDesc>
    <p>
<!-- Additional information -->
    </p>
  </profileDesc>
  <revisionDesc>
    <change>
<!-- Major stages for the electronic file -->
    </change>
  </revisionDesc>
</teiHeader>
```

The TEI supports two 'levels' or types of header

- *corpus level* metadata sets default properties for everything in a corpus
- *text level* metadata sets specific properties for one component text of a corpus

Corpus Header Example

```
<teiCorpus xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader type="corpus">
    <!-- corpus-level metadata here -->
  </teiHeader>
  <TEI>
    <teiHeader type="text">
      <!-- metadata specific to this text here -->
    </teiHeader>
    <text>
      <!-- ... -->
    </text>
  </TEI>
  <TEI>
    <teiHeader type="text">
      <!-- metadata specific to this text here -->
    </teiHeader>
    <text>
      <!-- ... -->
    </text>
  </TEI>
</teiCorpus>
```

Types of content in the TEI header

- free prose
 - prose description: series of paragraphs
 - phrase: character data, interspersed with phrase-level elements, but not paragraphs
- grouping elements: specialised elements recording some structured information
- declarations: Elements whose names end with the suffix Decl (e.g. subjectDecl, refsDecl) enclose information about specific encoding practices applied in the electronic text.
- descriptions: Elements whose names end with the suffix Desc (e.g. <settingDesc>, <projectDesc>) contain a prose description, possibly, but not necessarily, organised under some specific headings by suggested sub-elements.

A minimal header for 'Strange Meeting'

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>My 'Strange Meeting' document</title>
    </titleStmt>
    <publicationStmt>
      <p>An exercise for learning TEI.</p>
    </publicationStmt>
    <sourceDesc>
      <p> The primary resource of this file is
      <ref target="http://www.oucs.ox.ac.uk/wllit/collections/item/3350">
        Strange Meeting</ref> from Jon Stallworthy's edition, available on the First World War
        Poetry Digital Archive. </p>
    </sourceDesc>
  </fileDesc>
</teiHeader>
```

File Description <fileDesc>

- has some **mandatory** parts:
 - <titleStmt>: provides a title for the resource and any associated statements of responsibility
 - <sourceDesc>: documents the sources from which the encoded text derives (if any)
 - <publicationStmt>: documents how the encoded text is published or distributed
- and some **optional** ones:
 - <editionStmt>: yes, digital texts have editions too
 - <seriesStmt>: and they also fit into "series".
 - <extent>: how many floppy disks, gigabytes, files?
 - <notesStmt>: notes of various types

The File Description

- <titleStmt>: contains a mandatory <title> which identifies the electronic file (not its source!)
- optionally followed by additional titles, and by 'statements of responsibility', as appropriate, using <author>, <editor>, <sponsor>, <funder>, <principal> or the generic <respStmt>
- <publicationStmt>: may contain
 - plain text (e.g. to say the text is unpublished)
 - one or more <publisher>, <distributor>, <authority>, each followed by <pubPlace>, <address>, <availability>, <idno>

Title- and Responsibility- statements...

You can duplicate any of these elements as necessary, and document additional responsibilities with a generic **<respStmt>**:

```
<titleStmt>
  <title>Letter to Leslie Gunston</title>
  <author>Wilfred Owen</author>
  <editor>Renée van Baalen</editor>
  <principal>James Cummings</principal>
  <meeting>Introduction to TEI Course</meeting>
  <respStmt>
    <resp>Improved encoding</resp>
    <name>James Cummings</name>
  </respStmt>
</titleStmt>
```

<editionStmt> and <extent> example

<editionStmt> can be used to document the details of this particular edition:

```
<editionStmt>  
  <edition>First Edition</edition>  
</editionStmt>  
<extent>6.5 kb</extent>
```

<extent> is sometimes used to document number of words in a corpus.

<publicationStmt> example

```
<publicationStmt>
  <publisher>TEI @ Oxford</publisher>
  <pubPlace>
    <address>
      <orgName>IT Services, University of Oxford</orgName>
      <street>13 Banbury Road</street>
      <settlement>Oxford</settlement>
      <postCode>OX2 6NN</postCode>
      <country>United Kingdom</country>
    </address>
  </pubPlace>
  <distributor>IT Services, University of Oxford</distributor>
  <authority>James Cummings</authority>
  <date when="2014-10-10">10 October 2014</date>
  <idno>tei-ex03</idno>
  <availability>
    <licence>Licensed with a <ref target="http://creativecommons.org/licenses/by/3.0/">Creative
      Commons Attribution</ref> licence.</licence>
  </availability>
</publicationStmt>
```


<notesStmt> example

<notesStmt> can contain notes on almost any aspect of the file or its contents:

```
<notesStmt>  
  <note>Transcribed for TEI Workshop</note>  
</notesStmt>
```

These notes can be short statements, or many paragraphs long.

The Source Description

All electronic works need to document their source, **even** 'born digital' ones! There are variety of ways to do so:

- prose description, just a <p>
- <bibl> : contains free text and/or any mixture of bibliographic elements such as <author>, <publisher> etc.
- <biblStruct> contains similar elements but constrained in various ways according to bibliographic standards
- <biblFull> special-cases texts which were born TEI by replicating an embedded <fileDesc>
- A <listBibl> may be used for lists of such descriptions, e.g. bibliographies
- Specialised elements for spoken texts (<recordingStmt> etc.) and for manuscripts (<msDesc>) **Discussed later!**
- Authority lists: <listPerson>, <listPlace>, <listOrg>

<sourceDesc> example

```

<sourceDesc>
  <biblStruct>
    <analytic>
      <title>Letter to Leslie Gunston</title>
      <author>Wilfred Owen</author>
    </analytic>
    <monogr>
      <title>The Wilfred Owen Collection</title>
      <ref target="http://www.oucs.ox.ac.uk/wwllit/ collections/document/5243/4769"> First World War
Poetry Digital Archive</ref>
      <imprint>
        <publisher>The First World War Poetry Digital Archive</publisher>
        <pubPlace>Oxford</pubPlace>
        <biblScope type="pp" n="2">Two pages</biblScope>
      </imprint>
    </monogr>
    <relatedItem>
      <bibl>The source of this digital resource is a copy from the <istributor>Harry Ransom
Centre</istributor>.</bibl>
    </relatedItem>
  </biblStruct>
</sourceDesc>

```

<bibl> vs. <biblStruct> Example

```
<bibl>
  <title level="a">Enigma</title>, in
  <title level="j">Punch: or the London Charivari</title> (July 1, 1914), vol
  147, pp. 1-20
</bibl>
```

```
<biblStruct>
  <analytic>
    <title level="a">Enigma</title>
  </analytic>
  <monogr>
    <title level="j">Punch: or the London Charivari</title>
    <imprint>
      <pubPlace>London</pubPlace>
      <date when="1914-07-01">July 1, 1914</date>
      <biblScope type="vol">147</biblScope>
      <biblScope type="pp">1-20</biblScope>
    </imprint>
  </monogr>
</biblStruct>
```

Association between header and text

By default everything asserted by a header is true of the text to which it is prefixed. This can be over-ridden:

- as when a text header over-rides or amplifies a corpus-header setting
- when model.declarable elements are selected by means of the *@decls* attribute (available on all model.declaring elements)
- using special purpose selection/definition elements e.g. <catRef> and <taxonomy>

Most components of the encoding description are declarable.

Header: <encodingDesc>

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <!-- title/author/editor information -->
    </titleStmt>
    <publicationStmt>
      <p>
        <!-- Publication Information -->
      </p>
    </publicationStmt>
    <sourceDesc>
      <p>
        <!-- Metadata about the source -->
      </p>
    </sourceDesc>
  </fileDesc>
  <encodingDesc>
    <p>
      <!-- Description of the encoding -->
    </p>
  </encodingDesc>
  <profileDesc>
    <p>
      <!-- Additional information -->
    </p>
  </profileDesc>
  <revisionDesc>
    <change>
      <!-- Major stages for the electronic file -->
    </change>
  </revisionDesc>
</teiHeader>
```

Encoding Description

<encodingDesc> groups notes about the procedures used when the text was encoded, either summarised in prose or within specific elements such as

- <projectDesc>: goals of the project
- <samplingDecl>: sampling principles
- <editorialDecl>: editorial principals, e.g. <correction>, <normalization>, <quotation>, <hyphenation>, <segmentation>, <interpretation>
- <classDecl>: classification system/s used
- <tagsDecl>: specifics about usage of particular elements

Detailed notes in <encodingDesc> could be used to generate section of an editorial description.

<encodingDesc> Example (1)

```
<encodingDesc>
  <projectDesc>
    <p>The TEI@Oxford project created these teaching materials for teaching TEI workshops.</p>
  </projectDesc>
  <editorialDecl>
    <correction>
      <p>Apparent errors have been marked as <gi>sic</gi> but correct
        readings not provided.</p>
    </correction>
    <hyphenation>
      <p>Hyphens have been transcribed as they appear.</p>
    </hyphenation>
  </editorialDecl>
</encodingDesc>
```


<encodingDesc> Example (2)

```
<encodingDesc>
  <classDecl>
    <taxonomy xml:id="part-of-speech">
      <category xml:id="adje">
        <catDesc>adjectives</catDesc>
        <category xml:id="AJ0">
          <catDesc>adjective (unmarked) (e.g. GOOD, OLD)</catDesc>
        </category>
        <category xml:id="AJC">
          <catDesc>comparative adjective (e.g. BETTER, OLDER)</catDesc>
        </category>
        <category xml:id="AJS">
          <catDesc>superlative adjective (e.g. BEST, OLDEST)</catDesc>
        </category>
      </category>
      <category xml:id="AT0">
        <catDesc>article (e.g. THE, A, AN)</catDesc>
      </category>
    </taxonomy>
  </classDecl>
</encodingDesc>
```

```
<w ana="#AJ0">brilliant</w>
```

<rendition> element

- **<rendition>**: structured information about appearance in the source document

```
<tagsDecl>
  <rendition xml:id="r-center"
    scheme="css">text-align: center;</rendition>
  <rendition xml:id="r-small"
    scheme="css">font-size: small;</rendition>
  <rendition xml:id="r-large"
    scheme="css">font-size: large;</rendition>
</tagsDecl>
```

which you can easily point to from the text:

```
<hi rendition="#r-center #r-large">this bit of text was large and centred</hi>
```

but compare:

```
<hi rend="large center">this bit of text was large and centred</hi>
```

<appInfo> element

<appInfo>: structured information about an application which has edited this TEI file

```
<appInfo>
  <application version="1.8.2.2"
    ident="IMT" notAfter="2014-06-01">
    <label>Image Markup Tool</label>
    <ptr target="#P1"/>
    <ptr target="#P2"/>
  </application>
</appInfo>
```

Header: <profileDesc>

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <!-- title/author/editor information -->
    </titleStmt>
    <publicationStmt>
      <p>
        <!-- Publication Information -->
      </p>
    </publicationStmt>
    <sourceDesc>
      <p>
        <!-- Metadata about the source -->
      </p>
    </sourceDesc>
  </fileDesc>
  <encodingDesc>
    <p>
      <!-- Description of the encoding -->
    </p>
  </encodingDesc>
  <profileDesc>
    <p>
      <!-- Additional information -->
    </p>
  </profileDesc>
  <revisionDesc>
    <change>
      <!-- Major stages for the electronic file -->
    </change>
  </revisionDesc>
</teiHeader>
```

Profile Description

A collection of descriptions, categorised only as ‘non-bibliographic’.

Default members of the model.profileDescPart class include:

- **<creation>**: information about the origination of the intellectual content of the text, e.g. time and place
- **<langUsage>**: information about languages, registers, writing systems etc used in the text
- **<textDesc>** and **<textClass>**: classifications applied to the text by means of a list of specified criteria or by means of a collection of pointers, respectively
- **<particDesc>** and **<settingDesc>**: information about the ‘participants’, either real or depicted, in the text
- **<handNotes>**: information about the hands identified in a manuscript

<creation> example

```
<creation>
  <date when="1918-05"/>
  <placeName>Ripon</placeName>
  <listChange ordered="true">
    <change xml:id="CHG-1">First stage, written in pencil in Owen's hand </change>
    <change xml:id="CHG-2">Second stage, revised in pencil in Owen's hand</change>
    <change xml:id="CHG-3">Fixation of the revised passages and further minor revisions by Owen using
ink</change>
    <change xml:id="CHG-4">Addition of another stanza with a different ink, probably at a later
stage</change>
  </listChange>
</creation>
```

Here <listChange> records stages in changes to the document. Further down, in <revisionDesc> the same element is used to record changes to the electronic file.

Language and character set usage

The `<langUsage>` element is provided to document usage of languages and writing systems in the text. Languages are identified by their ISO codes:

```
<langUsage>
  <language ident="en">English</language>
  <language ident="fr">French</language>
  <language ident="bg-cy">Bulgarian in Cyrillic characters </language>
  <language ident="bg">Romanized Bulgarian</language>
</langUsage>
```

Classification Methods

<textClass> provides a classification (by domain, medium, topic...) for the whole of a text expressed in one or more of the following ways:

using <catRef> direct reference to a locally defined (e.g. in the corpus header) category

using <classCode> reference to some commonly agreed and externally defined category

using <keywords> assign arbitrary descriptive terms taken from a bibliographic controlled vocabulary or a tag cloud

BNC Example

```
<profileDesc>
  <creation>
    <date when="1962"/>
  </creation>
  <textClass>
    <catRef target="#WRI #ALLTIM1 #ALLAVA2 #ALLTYP3 #WRIDOM5 #WRILEV2 #WRIMED1 #WRIPP5 #WRISAM3 #WRISTA2 #WRITAS0"/>
    <classCode scheme="DLEE">W nonAc: humanities arts</classCode>
    <keywords scheme="COPAC">
      <term>History, Modern - 19th century</term>
      <term>Capitalism - History - 19th century</term>
      <term>World, 1848-1875</term>
    </keywords>
  </textClass>
</profileDesc>
```

This categorization applies to the whole text. For more fine grained classification, use *@decls* on e.g. a `<div>` element to point to applicable variation in header.

Detailed characterization of a text

<textDesc> provides a description of a text in terms of its 'Situational parameters'

```
<textDesc n="novel">
  <channel mode="w">print; part issues</channel>
  <constitution type="single"/>
  <derivation type="original"/>
  <domain type="art"/>
  <factuality type="fiction"/>
  <interaction type="none"/>
  <preparedness type="prepared"/>
  <purpose type="entertain"
    degree="high"/>
  <purpose type="inform"
    degree="medium"/>
</textDesc>
```

<particDesc> example (1)

```
<particDesc xml:id="p2">
```

```
  <p>Female informant, well-educated, born in Shropshire UK, 12 Jan 1950, of  
  unknown occupation. Speaks French fluently. Socio-Economic status B2 in the PEP  
  classification scheme.</p>
```

```
</particDesc>
```

<particDesc> can just contain paragraphs of prose, or a more structured
<person> element in <listPerson>

<particDesc> example (2)

```
<particDesc>
  <listPerson>
    <person xml:id="SL">
      <persName>Stuart Lee</persName>
      <note>
        <ref target="http://users.ox.ac.uk/~stuart/Site/About_Me.html"> Stuart Lee's home page</ref>
      </note>
    <!-- We could give more details about Stuart here -->
    </person>
    <person xml:id="IH">
      <persName>Ian Hislop</persName>
      <birth>
        <date when="1960-07-13">13 July 1960</date>
        <placeName>Swansea, Wales, United Kingdom</placeName>
      </birth>
      <occupation>Satirist, Journalist, Presenter, Libel Defendant</occupation>
      <note>
        <ref target="http://en.wikipedia.org/wiki/Ian_Hislop"> Ian Hislop's entry in Wikipedia</ref>
      </note>
    </person>
  </listPerson>
</particDesc>
```

Header: <revisionDesc>

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <!-- title/author/editor information -->
    </titleStmt>
    <publicationStmt>
      <p>
        <!-- Publication Information -->
      </p>
    </publicationStmt>
    <sourceDesc>
      <p>
        <!-- Metadata about the source -->
      </p>
    </sourceDesc>
  </fileDesc>
  <encodingDesc>
    <p>
      <!-- Description of the encoding -->
    </p>
  </encodingDesc>
  <profileDesc>
    <p>
      <!-- Additional information -->
    </p>
  </profileDesc>
  <revisionDesc>
    <change>
      <!-- Major stages for the electronic file -->
    </change>
  </revisionDesc>
</teiHeader>
```

Revision Description

- A list of <change> elements, each with a @date and @who attributes, indicating significant stages in the evolution of a document. Most recent first.
- Can be grouped into <listChange> elements. Used here it is about the electronic file, used in <creation> it is about the document.
- Can be maintained manually, or done by means of a version control system (like SVN)

```
<revisionDesc>
  <listChange>
    <change when="2014-07-03">
      <persName>James Cummings</persName> improved the header.</change>
    <change when="2012-02">
      <persName>Renée van Baalen</persName> transcribed the <title>Letter to Leslie Gunston</title>
document. </change>
    </listChange>
  </revisionDesc>
```

Manuscript Description

Why are manuscripts special?

- Manuscripts are *unique objects*, sometimes of great cultural or political value.
- Books, by contrast, exist in multiple copies, and can be described adequately by well-established and formalised bibliographic conventions.
- For manuscripts, there are several traditions, often descriptive or *belle lettriste*, and little consensus.
- Manuscripts can be ancient or modern, and it benefits us to use the same forms of description for them.

Similar concerns apply to other text-bearing objects, and `<msDesc>` is often also used for incunabula, early printed books, or other non-manuscript objects with unique texts.

Book of Hours



Preface MS

Preface.

This book is not about heroes. English Poetry

is not yet fit to speak of them
deeds or

Is not yet fit to speak of
Now is it about battles ^{deeds or lands} and glory ^{of battles & lands}
^{about glory} might, unity, dominion or power
we are not about any ^{it} war.

except

about all the ~~poets~~ ^{poets} ~~are~~ ^{are} ~~concerned~~ ^{concerned} with Poetry.

The subject of War, and the pity of it.

The party is in the pity -

~~There is hesitation in~~ making public
budget, so

These Hypotheses are not for the construction of a model in no sense. It may be to

to this generation. They may be to the
next. If I thought the letter of the

to this generation.
next. If I thought the letter of the
book would last I would ^{not} have
used paper ^{any} more. The paper is warm, darker
used paper ^{any} more. The paper is warm, darker
used paper ^{any} more. The paper is warm, darker

used proper names.
That is why the Iran Pact must be truthful.

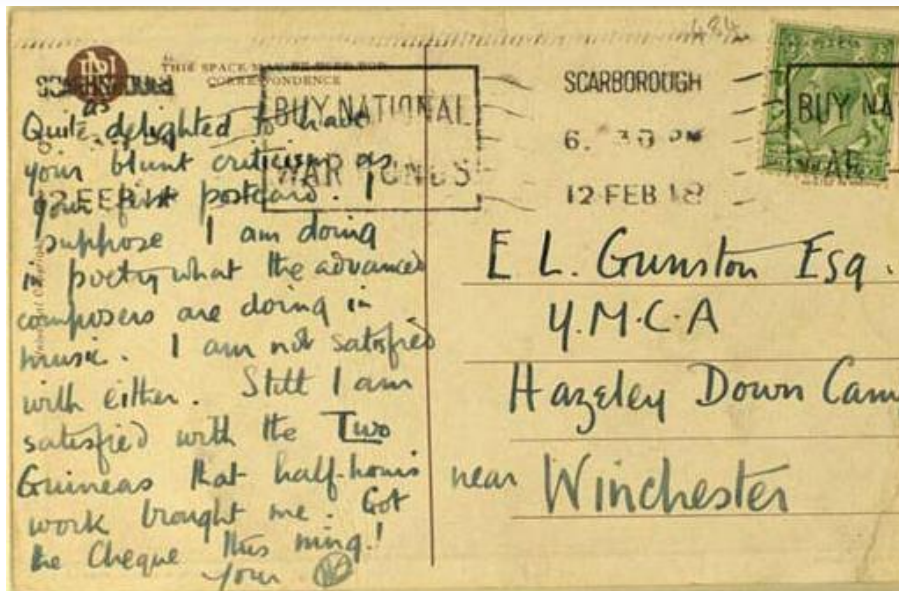
I thought the letters of the book were
 were might have used proper names; but if the spirit of
 it survives I am sure that the names will
 be content, for the future, to be
 proper fields, the flowers,
 for be, all of way, would be

Letter 1917-01-10 (page 5)

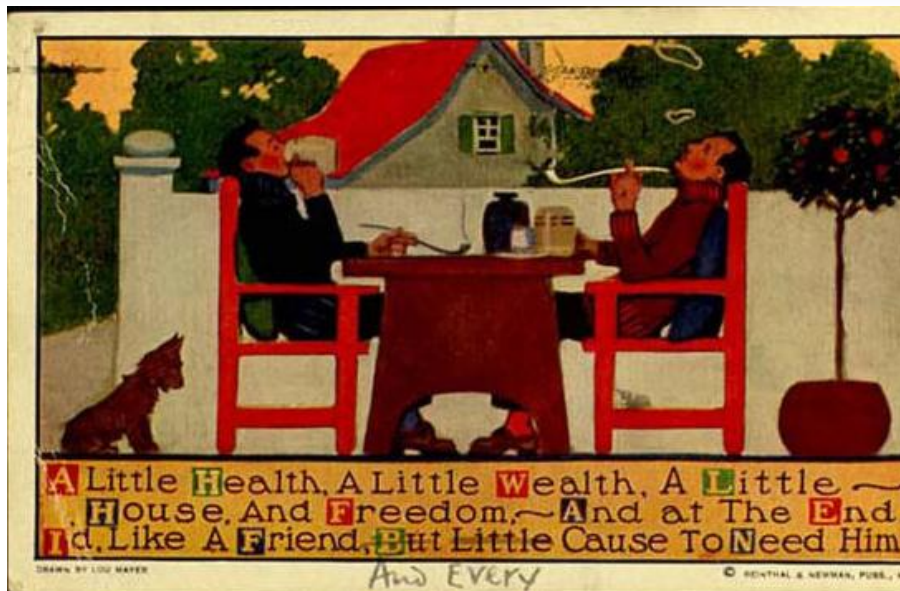
5/1
 I am not allowed to send a
 sketch, but you must know
 I am transformed now, wearing
 a steel helmet, buff jerkin of
 leather, & rubber-waders up
 to the hips, & gauntlets.
 But for the rifle we are
 exactly like Cromwellian
 Troopers. The waders are
 of course indispensable.
 In 2½ miles of trench which I
 waded yesterday there was
 not one inch of dry
 ground. There is a mean
 depth of 2 feet of water.
 It seems an era since
 Christmas, Day, and Goose,
 Carols, Dickens & Mistletoe.
 Assuming the war lasts another
 year I should get leave twice,
 or three times, for we get it,
 or should get it every 3 months.
 Be sure to have no chloride

Not so bad!

Postcard Back



Postcard Front



What about a sign?



What about a sign?



First Folio: MSND



Objectives of <msDesc>

The TEI <msDesc> element is intended for several different kinds of applications:

- standalone database of library records (*finding aid*)
- discursive text collecting many records (*catalogue raisonné*)
- metadata component within a digital surrogate (*electronic edition*)
- tool for 'quantitative codicology'

Catalogue Raisonné

An <msDesc> can appear anywhere a <p> paragraph can

```
<div>
  <head>The Arnamagnæan Manuscript Collection</head>
  <p>The Arnamagnæan Collection is widely recognised as one of the most
significant collections of early Scandinavian manuscripts in the world...</p>
  <p>Among its more important holdings are: <msDesc xml:id="AM02-0101"
    xml:lang="en">
<!-- ...-->
  </msDesc>
</p>
  <p>In the following manuscript... <msDesc xml:id="AM04-0595"
    xml:lang="en">
<!-- ...-->
  </msDesc>
</p>
</div>
```

Having one's cake and eating it

Two conflicting desires:

- preserve (or perpetuate) existing descriptive prose
- reliable search, retrieval, and analysis of data

The <msDesc> tries, wherever possible, to do both of these things.

Components of a manuscript description

Inside <msDesc> only <msIdentifier> is required. After this <head> is sometimes used for a brief summary. These are then followed either by one or more paragraphs (<p>), or one or more of the following specialised elements:

- <msContents>: alist of the intellectual content of the manuscript
- <physDesc>: groups information concerning all physical aspects of the manuscript
- <history>: provides information on the history of the manuscript, its origin, provenance and acquisition by current holding institution

Components of a manuscript description (cont.)

- **<additional>**: groups other information about the manuscript (e.g. administrative information relating to its availability, custodial history, surrogates)
- **<msPart>**: contains in essence a nested **<msDesc>**, in cases of composite manuscripts now regarded as constituting a single unit but made up of two or more parts which were originally physically distinct.

Structure of an <msDesc>

```
<msDesc xml:id="myFavouriteMS"
  xml:lang="en">
  <msIdentifier>
    <!-- Manuscript Identification -->
    </msIdentifier>
    <msContents>
      <!-- Intellectual Structure -->
      </msContents>
      <physDesc>
        <!-- Physical Description -->
        </physDesc>
        <history>
          <!-- Origin, Provenance, Acquisition -->
          </history>
          <additional>
            <!-- Additional Administrative Metadata -->
            </additional>
          </msDesc>
```

Inside the top level

Within each of these elements a number of sub-elements is available; `<msContents>`, for example, will normally consist of one or more `<msItem>` elements, each in turn containing specific elements for `<rubric>`, `<incipit>`, `<explicit>` and `<colophon>`, as well as the standard TEI elements `<author>`, `<title>` and `<bibl>` for bibliographic references. As with `<msDesc>` itself, however, the contents of these first-level and second-level elements need not be this structured, since there is also the option of just using paragraphs (`<p>`).

Identification (1)

The <msIdentifier>

Traditional three part specification:

- place (<country>, <region>, <settlement>)
- repository (<institution>, <repository>)
- identifier (<collection>, <idno>)

```
<msIdentifier>
  <country>United States of America</country>
  <region>Texas</region>
  <settlement>Austin</settlement>
  <institution> The University of Texas at Austin </institution>
  <repository>Harry Ransom Centre</repository>
  <collection>Wilfred Owen Collected Letters</collection>
  <idno type="folio">ff504</idno>
  <altIdentifier>
    <idno>Letter no. 535 Ed. 'Wilfred Owen Collected Letters'</idno>
  </altIdentifier>
  <msName>Letter to Leslie Gunston</msName>
</msIdentifier>
```

Identification (2)

```
<msIdentifier>  
  <country>Canada</country>  
  <settlement>Ottawa</settlement>  
  <repository>Library and Archives Canada</repository>  
  <collection>E.W.B. Morrison</collection>  
  <idno>MG 30 E 81 v. 16</idno>  
</msIdentifier>
```


Identification (3)

```
<msIdentifier>  
  <country>France</country>  
  <settlement>Troyes</settlement>  
  <repository>Bibliothèque Municipale</repository>  
  <idno>50</idno>  
</msIdentifier>
```

Identification (4)

Alternative or additional names can also be included:

```
<msIdentifier>
  <country>Danmark</country>
  <settlement>København</settlement>
  <repository> Det ArnamagnæanskeInstitut </repository>
  <idno>AM 45 fol.</idno>
  <msName xml:lang="la">Codex Frisianus</msName>
  <msName xml:lang="is">Fríssbók</msName>
</msIdentifier>
```

Structure of an <msDesc>: <msContents>

```
<msDesc xml:id="myFavouriteMS"
  xml:lang="en">
  <msIdentifier>
    <!-- Manuscript Identification -->
    </msIdentifier>
    <msContents>
      <!-- Intellectual Structure -->
      </msContents>
      <physDesc>
        <!-- Physical Description -->
        </physDesc>
        <history>
          <!-- Origin, Provenance, Acquisition -->
          </history>
          <additional>
            <!-- Additional Administrative Metadata -->
            </additional>
          </msDesc>
```

Intellectual Content

May simply use paragraphs of text or a set of `<msItem>` elements, optionally preceded by a prose summary. Unstructured:

```
<msContents>
  <p>An extraordinary charivari of heroic deeds and improving tales,
    including an early version of <title>Guy of Warwick</title> and
    several hymns.</p>
</msContents>
```

or structured:

```
<msContents>
  <summary>An extraordinary charivari of heroic deeds, improving
    tales, and hymns.</summary>
  <msItem>
<!-- details of Guy of Warwick here -->
  </msItem>
  <msItem>
<!-- other msItems for hymns here -->
  </msItem>
</msContents>
```

The <msItem> element

Manuscripts contain identifiable items, usually physically tied to a *locus*.

- <locus>, if present, must be given first
- then any of the following, in a specified order:
 - <author>, <respStmt>
 - <title>, <rubric>, <incipit>, <explicit>, <colophon>, <finalRubric>
 - <quote>, <textLang>, <decoNote>, <bibl>, <listBibl>, <note> ...
 - ... or indeed nested <msItem>s

<msContents> with multiple <msItem>s

```
<msContents>
  <!-- first item -->
    <msItem n="1">
      <locus from="5r" to="7v">fols. 5r-7v</locus>
      <title type="supplied">An ABC</title>
    </msItem>
  <!-- second item -->
    <msItem n="2">
      <locus from="7v" to="8v">fols. 7v-8v</locus>
      <title type="uniform" xml:lang="fr">L'envoy de Chaucer a Scogan</title>
    </msItem>
  <!-- ...further items here... -->
    <msItem n="6">
      <locus from="14r" to="126v">fols. 14r-126v</locus>
      <title type="uniform">Troilus and Criseyde</title>
      <note>Bk. 1:71-Bk. 5:1701, with additional losses due to mutilation throughout</note>
    </msItem>
</msContents>
```

Structure of an <msDesc>: <physDesc>

```
<msDesc xml:id="myFavouriteMS"
  xml:lang="en">
  <msIdentifier>
    <!-- Manuscript Identification -->
    </msIdentifier>
    <msContents>
      <!-- Intellectual Structure -->
      </msContents>
      <physDesc>
        <!-- Physical Description -->
        </physDesc>
        <history>
          <!-- Origin, Provenance, Acquisition -->
          </history>
          <additional>
            <!-- Additional Administrative Metadata -->
            </additional>
          </msDesc>
```

Physical Description

An artificial (but helpful) grouping of many distinct items.

You can simply supply paragraphs of prose, covering such topics as

- <objectDesc>: the physical carrier
- <handDesc>: what is carried on it
- <musicNotation>, <decoDesc>, <additions>
- <bindingDesc> and <sealDesc>
- <accMat>: accompanying material

Or, group your discussion within the specific elements mentioned above.

Similarly, within the specific elements, you can supply paragraphs of prose, or further specific elements.

The carrier 1

The <objectDesc> can contain just paragraphs, or <supportDesc> and <layoutDesc>

```
<objectDesc form="codex">
  <supportDesc material="mixed">
    <p>Early modern <material>parchment</material> and
    <material>paper</material>.</p>
  </supportDesc>
  <layoutDesc>
    <layout columns="1"
      ruledLines="25 32"/>
  </layoutDesc>
</objectDesc>
```

The carrier 2

A more complex substructure with specific elements for <support>, <extent>, <foliation>, <collation>, <condition>.

Multiple layouts may also be specified:

```
<layoutDesc>
  <layout ruledLines="25 32"
    columns="1">
    <p>
      <locus from="1r" to="202v"/>
      <locus from="210r" to="212v"/> Between 25 and 32 ruled
        lines.</p>
    </layout>
  <layout ruledLines="34 50"
    columns="1">
    <p>
      <locus from="203r" to="209v"/>Between 34 and 50 ruled
        lines.</p>
    </layout>
</layoutDesc>
```

<handDesc> and <decoDesc>

- <handNote> (note on hand) describes a particular style or hand distinguished within a manuscript.
- <decoNote> contains a note describing either a decorative component of a manuscript or a fairly homogenous class of such components.

<handDesc> example (1)

```
<handDesc hands="2">
```

```
  <p>The manuscript is written in two contemporary hands, otherwise unknown, but  
  clearly those of practised scribes. Hand 1 writes ff.1r-22v and hand 2 ff. 23  
  and 24. Some scholars, notably Verner Dahlerup and Hreinn Benediktsson, have  
  argued for a third hand on f. 24, but the evidence for this is insubstantial.</p>
```

```
</handDesc>
```

<handDesc> example (2)

```

<handDesc hands="2">
  <handNote xml:id="Eirsp-1"
    scope="minor" script="other">
    <p>The first part of the manuscript, <locus from="1v" to="72v:4">fols
1v-72v:4</locus>, is written in a practised Icelandic Gothic bookhand. This hand
is not found elsewhere.</p>
  </handNote>
  <handNote xml:id="Eirsp-2"
    scope="major" script="other">
    <p>The second part of the manuscript, <locus from="72v:4" to="194v">fols
72v:4-194</locus>, is written in a hand contemporary with the first; it can also
be found in a fragment of <title>Knýtlinga saga</title>, <ref>AM 20b II
fol.</ref>.</p>
  </handNote>
</handDesc>

```

Another <physDesc> example

```
<physDesc>
  <objectDesc form="folio">
    <supportDesc material="paper">
      <support>A single folio of <material>paper</material> in the collection as
ff504 recto and verso</support>
    </supportDesc>
    <layoutDesc>
      <layout columns="1"
        writtenLines="20">Written full width as a single column, with
approximately 20 lines per page</layout>
    </layoutDesc>
  </objectDesc>
  <handDesc hands="1">
    <handNote>Written in <persName ref="#W0">Wilfred Owen's</persName>
hand.</handNote>
  </handDesc>
</physDesc>
```

<additions>

The <additions> element can be used to list or describe any additions to the manuscript, such as marginalia, scribblings, doodles, etc., which are considered to be of interest or importance.

<additions>

<p>The text of this manuscript is not interpolated with sentences from Royal decrees promulgated in 1294, 1305 and 1314. In the margins, however, another somewhat later scribe has added the relevant paragraphs of these decrees, see pp. 8, 24, 44, 47 etc.</p>

<p>As a humorous gesture the scribe in one opening of the manuscript, pp. 36 and 37, has prolonged the lower stems of one letter f and five letters þ and has them drizzle down the margin.</p>

</additions>

<accMat>

<accMat> (accompanying material) contains details of any significant additional material which may be closely associated with the manuscript being described, such as non-contemporaneous documents or fragments bound in with the manuscript at some earlier historical period.

<accMat> A copy of a tax form from 1947 is included in the envelope with the letter. It is not catalogued separately. **</accMat>**

Structure of an <msDesc>: <history>

```
<msDesc xml:id="myFavouriteMS"
  xml:lang="en">
  <msIdentifier>
    <!-- Manuscript Identification -->
    </msIdentifier>
    <msContents>
      <!-- Intellectual Structure -->
      </msContents>
      <physDesc>
        <!-- Physical Description -->
        </physDesc>
        <history>
          <!-- Origin, Provenance, Acquisition -->
          </history>
          <additional>
            <!-- Additional Administrative Metadata -->
            </additional>
          </msDesc>
```

<history>

- <origin>: where it all began
- <provenance>: everything in between
- <acquisition>: how you acquired it

<origin> is *datable* element and thus has attributes *@notBefore* and *@notAfter*, *@when* etc.

<history> Example

```

<history>
  <origin>
    <p>Written in <origPlace>England</origPlace> in the <origDate notAfter="1300"
      notBefore="1200">13th cent. </origDate>
    </p>
  </origin>
  <provenance>
    <p>On fol. 54v very faint is <q>Iste liber est fratris guillelmi de buria de
    <gap reason="illegible"/> Roberti ordinis fratrum Pred<ex>icatorum</ex>
      </q>, 14th cent. (?): <q>hanauilla</q> is written at the foot of the page (15th cent.).</p>
  </provenance>
  <acquisition>
    <p>Bought from the Rev. <name type="person">W. D. Macray</name> on <date when="1863-03-17">March
      17, 1863</date>, for 1 pound 10s.</p>
  </acquisition>
</history>

```

Structure of an <msDesc>: <additional>

```
<msDesc xml:id="myFavouriteMS"
  xml:lang="en">
  <msIdentifier>
    <!-- Manuscript Identification -->
    </msIdentifier>
    <msContents>
      <!-- Intellectual Structure -->
      </msContents>
      <physDesc>
        <!-- Physical Description -->
        </physDesc>
        <history>
          <!-- Origin, Provenance, Acquisition -->
          </history>
          <additional>
            <!-- Additional Administrative Metadata -->
            </additional>
          </msDesc>
```

<additional> information

- <adminInfo> : administrative information
- <surrogates> : information about other surrogates, i.e. photographs, digital images etc.
- <accMat> : accompanying material
- <listBibl> : bibliography

Administrative information

- record history
- availability
- custodial history
- miscellaneous remarks

```
<adminInfo>
  <custodialHist>
    <custEvent type="conservation"
      notBefore="1961-03" notAfter="1963-02">
      <p>Conserved between March 1961 and February 1963 at Birgitte Dalls Konserveringsværksted.</p>
    </custEvent>
    <custEvent type="photography"
      notBefore="1988-05-01" notAfter="1988-05-30">
      <p>Photographed in May 1988 by AMI/FA.</p>
    </custEvent>
  </custodialHist>
</adminInfo>
```

And finally

A **<msDesc>** can contain **<msPart>**, essentially a nested **<msDesc>**, where originally distinct manuscripts or parts of a manuscripts have been brought together to form a composite manuscript.

```
<msDesc>
  <msIdentifier>
    <settlement>Amiens</settlement>
    <repository>Bibliothèque Municipale</repository>
    <idno>MS 3</idno>
    <msName>Maurdramnus Bible</msName>
  </msIdentifier>
  <!-- other elements here -->
  <msPart>
    <altIdentifier>
      <idno>MS 6</idno>
    </altIdentifier>
    <!-- other information specific to this part here -->
  </msPart>
  <!-- other msParts here -->
</msDesc>
```

Next

That is a whirlwind tour of the kinds of things you can put in the
<teiHeader> with some extra detail about <msDesc>!

Now let's do an exercise to remind ourselves