

Exercise 3: Creating a Manuscript Description

TEI @ Oxford

17 May 2013

1 Exercise 3: Creating a Manuscript Description

1.1 Learning Outcomes

When you successfully complete this exercise you should be able to:

- Use a TEI document you've not created
- Modify a basic manuscript description to provide more structure
- Understand the general categories of manuscript description
- Have experience editing a complex `<teiHeader>`

1.2 Summary

In this exercises you will add a manuscript description to an existing TEI document with an already detailed header. You'll modify an existing `<msDesc>` element with a basic structure to categorise manuscript description information into a more detailed structure. This will give you experience with a more complex document and creating a manuscript description in TEI.

1.3 Starting Up

Load up the file called 'letter.xml' from the materials directory. This file has been prepared for you as a starting point, and although it has a detailed `<teiHeader>` it is missing a manuscript description.

1.4 Reading through the file

This file contains a letter from Wilfred Owen to his cousin Leslie Gunston. It talks about a forthcoming address to the Field Club, and contains a partial draft of 'The Wrestlers'. It was written in July 1917 from Craiglockhart War Hospital, Edinburgh, Scotland. Images of this letter are available in the materials we've provided.

- Note the detailed `<teiHeader>` and what kind of information is recorded there.
- Look at the structure of the document as three divisions and make sure you understand these divisions.
- Note the use of the `<dateline>` element.
- See how the encoder has recorded line-breaks in the prose.
- What other elements has the encoder included? Make sure you understand the meaning of them. If you are unsure of the meaning of them, look them up on the TEI-C website at: <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/REF-ELEMENTS.html>

1.5 Inserting a basic `<msDesc>`

The information for our manuscript description will basically be taken from the document description at <http://www.oucs.ox.ac.uk/ww1lit/collections/document/5243>. But we'll pretend that we already have a basic manuscript description as part of a retrospective conversion project. There is no requirement with TEI `<msDesc>` to divide it into all the possible categories of information, instead all it requires is at least a `<msIdentifier>`, while any other information

1 EXERCISE 3: CREATING A MANUSCRIPT DESCRIPTION

could be stored in a few accompanying paragraphs. This is useful for the retrospective conversion of catalogues in other legacy formats to TEI XML where it is impossible to know that this paragraph is about the physical description or another one is about the provenance.

- Move the cursor to immediately following the closing `</listOrg>` tag and the comment saying to insert the `msDesc` file. At this point either cut and paste or insert (with the 'Document' -> 'File' -> 'Insert File') the file 'msDesc.xml' from the materials provided.
- As you'll notice, this contains a very basic `<msDesc>` with a minimal `<msIdentifier>`. This is all that is really required, but we can do better than that!

1.6 Filling out a `<msIdentifier>`

Let's expand the `<msIdentifier>`. As you have a lot more experience editing XML files in oXygen now, the steps will sometimes be given in less detail.

- Notice that the first paragraph mostly contains information that tells us where the manuscript is, in other words it identifies it and so this text could go in a `<msIdentifier>`.
- Take the information in this paragraph and expand the `<msIdentifier>` until it looks something like this:

```
<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>
```

- Note how elements are prescribed to appear in a particular order (from greatest level of granularity to more specific). Notice that most elements cannot be repeated (some like `<collection>` and `<altIdentifier>` can be).
- When you've finished creating the `<msIdentifier>` delete the remains of the first `<p>` from the basic manuscript description.

1.7 Providing some `<msContents>`

The second paragraph contains information that will be useful in compiling an `<msContents>`. This acts as a place to store structured information concerning the intellectual contents of a manuscript. It gives a place for a summary of the contents of the manuscript and multiple `<msItem>` elements form something like a table of contents of works in the document.

- Rename the second paragraph element as `<msContents>` (your document will now not be valid)
- Highlight the text inside from the start to the end of "Collected Letters.", press control-e to 'surround with element' and wrap this in a `<summary>`. This acts as a summary for the intellectual content

- Highlight the remaining text and surround it with a `<msItem>` element.
- Delete the 'Authored by' and surround 'Wilfred Owen (1893-1918),' with an `<author>` element.
- Surround 'English.' with a `<textLang>` element.
- Add an `@mainLang` with a value of 'en' (the ISO language code for 'English')
- Add a `@ref` to the `<author>` and point to your `<person>` for Wilfred Owen.
- As this `<msItem>` is recording information for this particular item we also want to give it a `<title>`. Create an empty `<title>` element and cut and paste "Letter To Leslie Gunston / The Wrestlers." into it.
- Your `<msContents>` should now look something like:

```

<msContents>
  <summary>"Letter To Leslie Gunston / The Wrestlers". Talks about forthcoming
address to the 'Field Club'. Includes a partial draft of 'The Wrestlers'. This
is letter no. 535 in Ed. 'Wilfred Owen Collected Letters'.</summary>
  <msItem>
    <author>Wilfred Owen (1893-1918).</author>
    <textLang mainLang="en"> English. </textLang>
  </msItem>
</msContents>

```

1.8 Giving a `<physDesc>`

The next paragraph has a lot of information about the physical aspects of the manuscript. Let's turn it into a `<physDesc>`

- Rename the `<p>` to be a `<physDesc>`
- Immediately inside this create an `<objectDesc>` with a `<supportDesc>` inside that.
- Inside that `<supportDesc>` add a `<support>`, and inside this put the text "A single folio of paper in the collection as ff504 recto and verso"
- You could wrap the element `<material>` around the word 'paper', but also you could add a `@material` attribute to `<supportDesc>` with a value of 'paper'.
- You could also categorise the object's form by adding a `@form` attribute on `<objectDesc>` with a value of 'folio'.
- After the closing `</supportDesc>` tag add a `<layoutDesc>` with a `<layout>` to record information about the physical layout. In this case "Written full width as a single column, with approximately 20 lines per page"
- To the `<layout>` element add a `@columns` attribute of '1', and a `@writtenLines` of '20'.
- After the closing `</objectDesc>` add a `<handDesc>` with a `@hands` attribute with a value of '1'.
- Inside the `<handDesc>` add a `<handNote>` with the remaining text "Written in Wilfred Owen's hand in pen.". You might want to mark Wilfred Owen as a `<persName>` with a `@ref` pointing back to the `<person>` for Wilfred Owen.

1 EXERCISE 3: CREATING A MANUSCRIPT DESCRIPTION

- Your `<physDesc>` now might look something like:

```
<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 in
    pen.</handNote>
  </handDesc>
</physDesc>
```

1.9 Detailing a `<history>`

The `<history>` element gives a place to detail the `<origin>`, `<provenance>`, and `<acquisition>` of the manuscript if available. In this case we have some minimal information about the origin of the manuscript

- Rename the second-last paragraph to a `<history>` element.
- Select all the text of "This letter was written by Wilfred Owen in July 1917 at Craiglockhart War Hospital." and surround it with a `<origin>` element.
- Inside this mark 'July 1917' as an `<origDate>` element. This is like the `<date>` element, but is specific to recording the origin date of the manuscript being described. Provide a `@when` attribute of '1917-07'.
- Similarly mark the 'Craiglockhart War Hospital' as an `<origPlace>` with a `@ref` of '#craiglockhart' to point to the `<place>` you made earlier. You could also surround the text with an `<orgName>` if you want to indicate that this is an organizational name. As before you could mark Wilfred Owen's name.
- Your `<history>` element should look something like:

```
<history>
  <origin>This letter was written by <persName ref="#W0">Wilfred Owen</persName>
  in
  <origDate when="1917-07">July 1917</origDate> at <origPlace ref="#craiglockhart">
    <orgName>Craiglockhart War Hospital</orgName>
  </origPlace>
</origin>
</history>
```

1.10 Noting `<additional>` Information

At the end of your `<msDesc>` you can include an `<additional>` element which stores other information such as `<adminInfo>` (for recording administrative events of the object), `<listBibl>` (for listing bibliographic citations about the object), and `<surrogates>` (for listing additional representations of the object).

- Change the final paragraph to an `<additional>` element with a `<surrogates>` inside that containing all the text.
- Modify the URL given to be a `<ptr>` with a `@target` attribute.
- Your `<additional>` element should look something like:

```
<additional>
  <surrogates>A digital image is available from the First World War Poetry Digital
  Archive
    at <ptr
      target="http://www.oucs.ox.ac.uk/ww1lit/collections/document/5243/4770"/>.
    </surrogates>
  </additional>
```

1.11 Saving Your Work

Let's save our work:

- Is your work well-formed? Do you have a happy green square or an angry red one?
- Have you formatted and indented your work automatically?
- From the 'File' menu select 'Save' or click on the Save icon (looks like a old-style 3.5" disk).
- Or if you prefer use the 'File' then 'Save As' menu item to save the file using the name 'exercise05.xml' or another name of your choice.

1.12 Differences between `tei_all` and the FIHRIST schema

The FIHRIST schema available from <http://www.fihrist.org.uk/about> has numerous differences from what is allowed in full TEI. This is derived from a subset customisation of the TEI Guidelines for manuscript description. Whereas the TEI allows many different forms of manuscript description, a large EU project called ENRICH (which we were major partners in) created a standardised subset of TEI primarily designed for cataloguing western European manuscripts. We were primarily responsible for wrangling the decisions of various manuscript cataloguers into a TEI customization. This schema was further modified by projects such as OCIMCO in the Bodleian to cater for describing Islamic manuscripts. The current FIHRIST schema dates from 2009, and a variety of things have changed in the TEI since then. In embarking on a new project it will make sense to re-examine the decisions made in the FIHRIST schema and see if they are still applicable.

If you use the menus 'Document' -> 'Schema' -> 'Associate Schema' and re-associate it with the FIHRIST schema 'enrich-wamcp.rnc' then you will find that oXygen will complain about many 'errors' in the document. These are mostly where content models have changed as the TEI has developed, or where the FIHRIST schema requires certain attributes, attribute values, or elements to be in a particular order/structure where TEI is more forgiving.

If we have time I'll go through these problems and explain what is required.

1.13 Self-Assessment

Check if you understand some of the core principles of this exercise by answering the following questions:

- What is the only required aspect of a TEI manuscript description?
- How does one record the separate works of intellectual content present in the manuscript?

1 EXERCISE 3: CREATING A MANUSCRIPT DESCRIPTION

- Where does one describe the support which forms the object, or its layout?
- How does one record the origin, provenance, and acquisition of the object?
- Where might you record the hands found in the document?