Editing and Encoding Instructions

## Preliminary Notes

The current Guidelines are version 4.0. I have attached a copy in Word format, and there is also a web-based version available at <http://scottkleinman.net/wp-content/uploads/guidelines-4.0/index>.

In the interest of time, we are starting with a template taken from Carl Horstmann’s edition of the *South English Legendary*. The HTML version produced for the *Middle English Dictionary* is available at <http://quod.lib.umich.edu/c/cme/AHA2708.0001.001/1:3.18?rgn=div2;view=fulltext>. The HTML has been converted to TEI/XML. In the life of Becket, some pre-processing scripts have been used to automatically insert highlighting of *litterae notabiliores* and of the first letter after a *punctus elevatus*. Since the scribe or rubricator does this regularly, in most cases the automatic highlighting will be correct. However, there are exceptions, so everything needs to be double checked against the manuscript.

## Accessing the Manuscript Images

An image viewer is available at <http://tsar192.grid.csun.edu/osdviewer.php>. Occasionally the server goes down, so please report the issue if you are not able to access the viewer. If this happens, there is a fallback scenario. Go to <http://scholarworks.csun.edu/handle/10211.3/123190> and search for the folio you are editing. Then click the **JPEG2000 image** link. For forther information, see the ImageAccessMethods.html file in myCSUNBox.

## Features of the XML File

1. Every AEME file has to have a <teiHeader>, a <facsimile>, and a <text> element. Since you will not be editing the <facsimile> element and will rarely need to edit the <teiHeader>, I have removed those to standoff files (facsimile.xml and teiHeader.xml). Line 4 includes those standoff files so that your code will validate in oXygen. **These files must be in the same folder as the file you are editing.**
2. The entities.dtd file (included in line 2) is also necessary for validation. It contains all the special character entity codes used regularly in encoding AEME editions. AEME uses the recommendations of the Medieval Unicode Font Initiative (MUFI) for the selection of special character entities. See the discussion in the Guidelines. For your convenience, a copy of the MUFI recommendations is included with this file. If you find yourself commonly using an entity that is not listed in entities.dtd, you can add it. However, make sure that you upload the new entities.dtd file back to the project repository (see **Submitting Your Work**). **The entities.dtd file must be in the same folder as the file you are editing.**
3. Line 7 begins a section in XML comments, meaning that anything you put between the <!--XXX --> will not be processed. You can use this section to keep a record of your work. At present, it contains a record of the pre-processing steps followed by a list of named entities in the text. See the next point about named entities.
4. Named entities are people and places, coded as <name type="person" ref="#bioXXX"> or <name type="place" ref="#locXXX">. By convention, I replace XXX with the name **in Modern English spelling** followed by 0001 for people and 01 for places. It would be helpful if you stick to this convention. Two examples have been placed at the top of the St Mark file to demonstrate the format. You may also add information such as dates or other explanatory material to help identify historical figures.
5. Each folio begins with an XML comment with the folio number, e.g. <-- folio 61 -->. This is merely for easy navigation. The numbering is taken from Horstmann; if there is no letter after the number, the folio is a recto. If the letter "b" follows the number, the folio is a verso. Underneath the comment is the actual AEME markup (e.g. <pb n="61r">) followed by the column number (<cb n="1"/>). You should not need to modify this code, except in one respect. The conversion of Horstmann’s edition from HTML to XML placed the folio division on the wrong line. You may need to copy and paste the comment and the <pb> and <cb> one line above its current location.
6. All lines begin with <hi rend="touched-red">X</hi><sg type="ln"/>, since this is the regular pattern. However, you will need to check that it is correct in every instance.
7. Where the line begins with &para;, the following code is used:

<hi rend="h0 red-ink">

You will need to change h0 to the correct height change red-ink to blue-ink in the individual instance. Also, check the surrounding code to make sure it accurately reflects the manuscript.

1. The first letter following a *punctus elevatus* has been automatically highlighted in red. In some cases, this may not reflect the manuscript reading, so check every line. Also, some lines may have highlighting of other letters, and you will need to add the code for that.
2. Long *s* regularly occurs before t and round r regularly occurs after *o*, *b*, *p*, and *d*. These have therefore been added automatically. You will need to check readings for accuracy as well add long *s* and round *r* where the pre-processing has not captured the manuscript reading.
3. All the punctuation that is not enclosed in <pc> is Horstmann’s modern punctuation and should be removed.
4. Horstmann’s notes are set off by <note type="enc" ref="#bioHO0001">. You will want to remove these notes and replace them with appropriate markup or notes of your own.
5. Some of Horstmann’s emendations appear in square brackets. You will want to replace that with appropriate markup. For instance, where you see "fol[c]", this should be changed to

<choice>  
 <sic>fol</sic>  
 <corr>folc</corr>  
</choice>

If you are not sure you agree with Horstmann’s emendation, you can attribute it to him as follows:

<corr resp="#bioHO0001">folc</corr>

1. Horstmann generally follows the manuscript capitalisation, but you should use <choice> to adjust it to modern usage in the critical markup. See the Guidelines for discussion.
2. Finally, there are all the usual idiosyncratic things you come across in the manuscript: corrections, damage, additions, and so on. It will be your task to add the markup for these features as you encounter them. If you are unsure how to handle individual features, and the Guidelines are unhelpful, please ask for help.

## Summary of Editing/Encoding Tasks

1. Ensure that all named entities are correctly encoded and recorded in the comments at the top of the file.
2. Ensure that folio divisions are in the right place.
3. Check all highlighting and add appropriate markup where necessary.
4. Correct markup of ornamental capitals, where appropriate.
5. Convert *s* and *r* to long *s* and round *r* where appropriate. In some cases, you may need to correct conversions done automatically.
6. Remove Horstmann’s modern punctuation.
7. Remove Horstmann’s notes and replace them with appropriate markup or your own notes.
8. Remove Horstmann’s emendations in square brackets and replace them with appropriate markup.
9. Use <choice> to encode correct modern capitalization in the critical markup.
10. Encode any features you observe in the manuscript that are not covered by the above tasks.

## Submitting Your Work

1. You should **submit (by e-mail)** a weekly record of hours spent and tasks accomplished so that I can track your work for payment purposes.
2. The project maintains a repository on GitHub at <https://github.com/scottkleinman/aeme/tree/master/bodllaudmisc108>. There you can find many samples of code (in various stages of the editing process). If you are comfortable using GitHub, I can make you a collaborator for the repository, and you can push your work directly to GitHub. If not, you will have to e-mail me your completed work (the xml file of the text and the entities.dtd file, if you have modified it). Alternatively, I can make you a member of my students’ group folder, and you can submit work that way. Let me know which method you would prefer.

## Proofreading

We have a method of generating HTML output for TEI files, which makes proofreading much easier. I will send you details of this method at a later date.

## Files Included

* 3ap\_SEL-Mark.xml
* entities.dtd
* teiHeader.xml
* facsimile.xml
* Guidelines 4.0.docx
* MUFI-Alphabetic-3-0.pdf