Mikael Gunnarsson at Wed, 19 Oct 2022 17:53at Wed, 19 Oct 2022 17:53

Jacqueline, Elisabeth I have read your report, which is sufficient and actually a bit more than that with respect to documentation. It still remains a bit unclear how you think in terms of long-time preservation. You have sent me a huge archive with lots of images in different forms among other things and you plan for a publication when copyright issues are clarified, right? Does it involve more duplication? There actually seem to be many files for each scanned source. I sense some systematics in this, but I wonder if on one hand all these jpegs really are needed and on the other hand if you yourselves in a couple of years are able to determine the systematics in the file archive. Would it not be enough to preserve only the TIFF files and furnish them with embedded metadata, which is lacking right now?

A few comments on your report. Do not underline or color URLs in the bibliography. You are listing 4 chapters in A Companion to digital humanities but it is unclear why you do that and if you have been using them. The TEI guidelines are like the RDA Toolkit. Do not list it in the bibliography if you are not quoting from it and in that case be precise It is not easy to identify to which image the captions pertain, the previous one or the succeeding one. You say that you did try the Book scanner but couldn't get anything but PDF from it. This one says otherwise https://www.ejmotiwalla.com/Codes/Maxphotos/Geomatictools/Previews/bookeye2-plus.htm Anyhow you have did a good work on scanning. You are saying "Before encoding began, decisions were made regarding what types of information should be described, how it should be described, utilising the TEI namespace, and how to arrange the text to reflect the format of the original document". However, you have the look and feel of the source as images, so you really do not need to bother about imitating the "format" of the source, if you mean the location and placement of text snippets. You have otherwise done a meticulous good and ambitious work on the TEI/XML encoding.

However, you do not clarify how you did go from TEI/XML to HTML. You are only saying you had problems and it is allowed to do some post HTML editing. I for once cannot altogether map the contents of the TEI file with the contents of the html files.

Here is the main problem. You have completely messed up the file paths so that neither do style sheets or images show up when trying to view the html files in a browser. I tried to fix that for the gallery.html but that resulted in images that were too large to fit on the screen, they probably need the width attribute. As a result I have not had the opportunity to assess your full work, and even though the images and the TEI seems very good, you need to come back with these issues fixed.

There are basically two different ways to go about with this. Either you rearrange the file structure to match the expected structure from the TEI and HTML files (but I assume you need to revise some XPaths as well) or you correct the XPaths in the XSL to reflect the existing file structure. Finally, if you fix this issue I would not go for the lowest grade (E). It is a lot better than that.

**Check List :**

MUST DO

* ~~Fix file paths for stylesheet and images (in HTML and TEI)~~
* Either you rearrange the file structure to match the expected structure from the TEI and HTML files (but I assume you need to revise some XPaths as well) or you correct the XPaths in the XSL to reflect the existing file structure.
* ‘cannot altogether map the contents of the TEI file with the contents of the html files. ‘

COMMENTS

RE: Filing + Pics

* Systematic filing – in support of preservation
* JPEGs needed?
* Embed TIFFs with metadata

Report

* Comments on long term preservation in the report
* ~~Do not underline coloured links~~
* ~~Remove chapter listing and TEI from references list~~
* ~~Cannot identify which imgs match with the captions (Jackie)~~
* Justify book scanner
* Did not mention ‘how you did go from TEI/XML to HTML’