

LEAP Transcription Kick Off – Documentation (Revised)

Adrian S. Wisnicki (28 May 2014)

Overview

This document sets out the procedure for transcription for LEAP. It reviews the information presented in the LEAP transcriptions pre-kickoff (25 April 2014) and revises (and considerably simplifies) that procedure to reflect lessons learned in our transcription test phase (26 April – 27 May 2014).

To carry out transcription for the project, four elements are needed: 1. The *Livingstone Online* Encoding Manual (LEM), 2. The appropriate LEAP TEI P5 template, 3. The LEAP RelaxNG Schema (leap.rng), and 4. Images of the item to be transcribed.

1) The *Livingstone Online* Encoding Manual (LEM)

The *Livingstone Online* Encoding Manual (LEM) is currently found at the following URL:

<http://research.it.ox.ac.uk/LEAP/> The LEM can be used online or downloaded and used locally. To download either select the “ZIP of all the above” option from the URL above or right click on the LEM, select “Save As” and save it to your hard drive. If you use the LEM locally, you should also download the leap-doc.css file from the URL above in the same way as the LEM and place it in the same directory as the LEM. The css file provides styling for the LEM.

The LEM was developed in collaboration with James Cummings and represents a revised, enhanced, and expanded version of the original *Livingstone Online* cheat sheet. The LEM consists of two principal parts: 1) The numbered opening section and 2) the appendices. The numbered opening section (about 35 pages when printed) is self-explanatory, is what you will normally use for transcriptions, and will cover the majority of encoding issues encountered.. The appendices supplement and expand upon the numbered section and can be ignored unless needed. The LEM will eventually be incorporated into the new *Livingstone Online* site.

2) The LEAP TEI P5 Transcription Templates [updated]

Henceforth, for each diary that we transcribe, I will provide a diary-specific template, for instance, liv_000002-template (see point #4a, below). This diary-specific template will require you to complete two distinct sections: the <teiHeader> and the <text> sections. The <teiHeader> section will be mainly completed; you will need to change only a few bits (that will be marked as needing change). The <text> section will consist of your transcription of the given Livingstone document based on the LEM.

3) The LEAP RelaxNG Schema [updated]

The LEAP RelaxNG Schema (leap.rng) is also available for download from the URL given in point #1 above. This schema will allow us to validate our transcriptions against our new *Livingstone Online*/LEAP TEI specifications.

Download this schema and place it in the same folder where you keep your transcriptions. Validation and association will be automatic when you open your transcriptions in Oxygen.

4) Images of the Item to be Transcribed

For LEAP the principal objects that we are transcribing are Livingstone's Field Diaries and the Unyanyembe Journal. The University of Glasgow is creating images of these items (as well as many, many other items from the David Livingstone Centre) for our project.

a) Accessing the Images and Diary Template [updated]

To access these images, navigate to the following directory:

[removed]

In this directory you will find two kinds of folders: 1) folders that contain images for a given diary and 2) folders in which you will put your transcriptions of the given diary.

The folder names represent enhanced versions of the LEAP file names, so as to make identification easier. For Field Diary II (liv_000002), for instance, you will find the following two folders:

liv_000002-Field-Diary-II-images

liv_000002-Field-Diary-II-transcriptions

The transcriptions folder will always contain the template for the given diary (see point #2, above).

Once you have found the images of the segment of the diary that has been assigned to you, download them onto your computer. Transcription assignments will be made by Megan and coordinated via a spreadsheet in Google docs. Please check with Megan if she has not "shared" this spreadsheet with you.

b) Naming the File of Your Transcription

Section 2.3 of the LEM (<http://research.it.ox.ac.uk/LEAP/leap.html#file-naming>) contains our general file naming rules for *Livingstone Online*. For LEAP, each of us will always be transcribing only a segment of a given item. As a result, please review the rules in the LEM and follow them, but please take note that when you name your transcription for LEAP, you will also be using the page number segment in the file name, and this page number segment should correspond to the first image of the file that you are transcribing.

So for instance, if I were transcribing pages 10 to 20 of Field Diary II, the name of the XML file containing transcriptions of all these pages after my initial transcription would be "liv_000002_0010_aw.xml". After it had been proofread by Megan Ward it would be "liv_000002_0010_aw_mw.xml". Once all the separate transcribed parts of this diary were assembled into the final integrated file, the file would be called "liv_000002.xml".