

## Open Philology Editing Environment (OPEn) User Instructions

This is a simple introduction to OPEn for first time users. It will help you get started with your first editing Project.

The Introduction does not get into any detail about the editing process itself: once you have started your Project, you will find that the different functionalities *within* the app are on the whole self-explanatory. The challenge is to get a new Project going, and this is what this Instruction will help you do.

OPEn was designed to be compatible with CollateX output files. In principle, while it is possible to start with a file in any format (including plain text), the key functionalities of OPEn will work only if the file is 'tokenized' ideally in .json format.

If you have a single witness of your text to start with, perhaps in a plain text format, find a way to tokenize that file and convert it into .json.

If you have multiple witness files, we recommend that you collate them using the CollateX script.

### **Pre-processing**

But before you run CollateX on your files, we strongly suggest that you pre-process your files by cleaning them up of unwanted signs and by standardizing their format. This will greatly facilitate your editing work later on.

Eventually we would like to write a script to do this pre-processing automatically, but for the time being it is not too difficult to do it by hand.

Below we outline the procedure that we like for pre-processing of witnesses of Tibetan texts in romanized Wylie transliteration. (For e.g. Chinese, the procedure will be slightly different).

The guidelines below include specific instructions for the use of the 'replace' function with regular expressions (RegEx). For this, you can use any program whose 'replace' functionalities allow the application of RegEx. There are many such programs available.

In each of your witness files individually:

1. Replace backslashes with vertical bars
2. Replace double shads with space in between with double shads with no space in between
3. Add space after last letter and before single shad  
Place in the 'find' line: **([a-z])(\|)**  
Place in the 'replace' line [NB: there is a space between 1 and \| in the replace line, this is the space you are adding!]: **\1 \|2**
4. Add space after shad and before first letter  
'Find' line: **(\|)([a-z])**  
'Replace' line: **\1 \|2**
5. Add space after shad and before a'chung  
'Find' line: **(\|)(')**  
'Replcae' line: **\1 \|2**

## Run CollateX

Once you have pre-processed your witness files, it's time to run CollateX on them.

Download CollateX from <https://collatex.net/>

Run the collatex script with the following arguments:

-f json

-t

The basic command will look like this:

```
java -jar collatex-tools-1.7.1.jar -f json -t -o [outputfilename.json] [file1.txt] [file2.txt] [file3.txt]...
```

Ideally you want to keep your input witness files in the same folder in which you have the CollateX script.

Depending on the size and number of your input witness files, the CollateX operation can take a long time, so be patient!

The output file (in this case, 'outputfilename.json') is the file you will then feed into OPEn when you 'Start a New Project.'

### Add witness sigla to json file

But before you can import the json file into OPEn to start your New Project, you need to do open the json file in any text editing program (e.g. Notepad on a Windows machine or Textedit on a Mac) and manually insert the witness sigla corresponding to the input witness files that you collated in CollateX.

For example, if in the CollateX operation above file1.txt contained the text of the Basgo, the corresponding siglum will be 'Ba.' If file2.txt was the Derge woodblock print, the siglum will be 'D.' And so on. For sigla of Tibetan canonical collections, we follow the guidelines recommended by Paul Harrison and Helmut Eimer, "Kanjur and Tanjur Sigla: A Proposal for Standardization," in *Transmission of the Tibetan Canon: Papers Presented at a Panel of the 7th Seminar of the International Association for Tibetan Studies*, Graz 1995, edited by Helmut Eimer (Verlag der Österreichischen Akademie der Wissenschaften, 1997).

The result you are aiming at is a .json file that begins with:

```
{"witnesses":["Siglum for file1.txt","Siglum for file2.txt","Siglum for file3.txt"...],"table":...
```

For example:



```
File Edit View
{"witnesses":["Ba","D","Go","He","J","V"],"table":
"},[["di "],[["dang "],[["skad "],[["skad "],[["
"],[["ye "],[["gis "],[["gis "],[["
```

### Start your New Project

Now you're ready to Start New Project in the app. Simply import your .json file and start editing.