

Topic: Coordinating SARIT development between New York and Heidelberg

BK = Birgit Kellner, PMA = Patrick McAlister, PO = Patrick Olivelle, AO = Andrew Ollett, JP = Jens Petersen, DW = Dominik Wujastyk

1. New online interface: See JP's recent [communication](#) regarding standoff markup. Questions about the development of an online editor for TEI files using standoff markup:

- AO has a number of questions about the markup of the base texts (what markup is inline in the base text, and what markup is standoff?). These questions should be addressed in the next meeting (October). Because there are ways of transforming inline markup to standoff markup, the CU team will proceed with processing texts until then. JP will put examples of standoff markup in our shared Dropbox folder.
- DW asks whether standoff markup might introduce problems as in the old Microsoft Word format (where formatting information etc. was linked to the base text by bytemarks). JP says that in the proposed system, the markup is indexed to the base text by xml:id and not by bytemark.
- PO asks whether the software will allow people who are not familiar with TEI/XML to edit the texts. (Yes, says JP.)
- Who will pay to develop the software? JP: HRA might take it on as one of their own projects. We will know whether the HRA has greenlighted the project in about a month (after the next budget meeting). BK: The HD subproject has budgeted some DFG funds for programming which could be used for this purpose. BK & JP: The HRA may be interested in developing this software even if it were not required by the SARIT project, because there is nothing like it currently available and there is a lot of interest in a collaborative editor for TEI projects.
- What kinds of TEI files will this software use? BK & JP: Base text may have very simple (structural) TEI encoding that is already present in files from SWIFT. AO: Some texts in CU subproject already have more complex encoding (e.g., *Abhinavabhāratī*); we should look at these files and see if anything will be problematic for the proposed editor.

2. Sharing data:

- For TEI files:
 - Dropbox: It is convenient and suitable for people who are not familiar with version-control. It is not good with version control, but in our project, one person usually works on a file at a time. DW adds that Dropbox has an extra service that keeps track of all changes (not just those made within the past month). DW suggests getting a corporate or university Dropbox account (around \$700). We may be able to make do with our current (personal) accounts. AO will share the CU subproject's Dropbox folder with the group.
 - Git: JP raised some questions about using Git with XML files. PMA uses it on several projects and says that it usually manages to merge even scripts and XML files successfully. PMA notes that files on the SARIT project on git are not necessarily loaded onto Philologic so the same project could host the "development" files (those currently in processing) and the "finished" files. AO

mentions that the "development" files should be private because some of the texts in processing are still in copyright. PMA confirms that the file-change history can be transferred from one Git project to another (DW's question).

- Dropbox & Git (PMA's suggestion): If each subproject has its own Dropbox directory, RAs could edit the shared files and the coordinator could then copy the files to another directory on his/her computer that is hooked into a Git directory. That way, people who aren't familiar with Git don't have to deal with it, and the coordinator can still take advantage of the way that Git merges files and tracks changes.
- Software: This will be available, at least to project members, over the course of development.
- Databases: AO is working on the SKSEC database.

3. Schedule:

- Synchronizing schedules: Each subproject (CU and HD) will have its own schedule with milestones. BK suggests that AO draw up a provisional schedule for the CU subproject to be shared before our October meeting, so that the HD team and the HRA can build their own schedules off of it. AO will put this schedule on the wiki.
- Conferences: We will plan to present our own on SARIT at a number of conferences.
 - TEI Conference. Submission deadline for 2014 (16th to 19th February in Munich) is already August 1 2013. We should plan to present at the 2015 conference. (<http://www.tei-conf.org/14/papers.php>)
 - IABS Conference Vienna 2014 (second last week of August - panel submission deadline already past, BK will be there)
 - Fifth International Dharmakīrti Conference in Heidelberg 26-31 August 2014 organized by Buddhist Studies in Heidelberg -> presentation should be envisaged, but a meeting of all involved?
 - World Sanskrit Conference Bangkok: 28 June - 02 July 2015: We should submit a panel, and talk to Oliver Hellwig and Amba Kulkarni to see what they are planning to present and whether they would participate in a SARIT panel.
- Announcements: DW suggests that we announce the project with a short article in an open-access journal, like the TEI journal. BK cautions against announcing anything too soon. The HRA's project, an online editor that uses standoff markup, should be announced in the TEI journal when it's approved and begun (in several months). The Indological aspects of the project could be discussed in a short article as well. AO notes that we should have an accessible set of guidelines (DW: "How you can contribute") for scholars who want to participate in the project, but this will have to wait several months until we come to a consensus about markup.

4. Project communication: As it stands, we're sharing information in three ways:

- The sarit-dev wiki: anything that we'll want to read more than once, and anything that requires discussion, should be posted here. Schedules and meeting agendas will be posted here as well.
- The sarit-development email list: Announcements (including of major updates to the wiki) should be posted here, but longer discussions---especially of a technical nature---should take place on the wiki. BK notes that some subscribers to sarit-development may not want to receive all our minutia. We could create a separate list, but it seems better at this point to let subscribers deal with this (e.g. by managing their subscription settings).

- Dropbox: this will mostly be used for our text files.
- 5. The **next meeting** will take place in October. The focus will be on TEI markup, and it will include the new project coordinator at Heidelberg. The Gandhari project of Munich will be visiting Heidelberg and it might make sense to include them in our meeting. Everyone will check their schedules and we will create a Doodle in due time. DW suggests that we should think of the project as providing a "killer app"---something that, in a few years, scholars who work on Sanskrit texts wouldn't think to do without.
 - Points for future discussion: (1) Should we strive to implement the **same encoding** in all SARIT texts? (This is not trivial: our encoding is so far based on our source-texts, and our source-texts vary greatly in the way they report variants, etc.) (2) **How** should we deal with particular encoding issues (e.g., how to encode textual variation—the TEI P5 allows one to encode variation in several ways, but presumably we will need *one* way that will work with our online interface)? [NB Parellel segmentation is the only format currently supported by Juxta, so I think we should use that -- DW] Should we discuss/document them [on the wiki](#) or elsewhere? (3) **When** should we have dealt with all of these encoding issues? Will we or any of our partners need a “complete” set of guidelines, or even a SARIT **schema**?