**Transcription meeting 2/26/14**

Present: William, John, Jenna, Joe, Mike, Trish

Absent: Francis, Joel

Agenda:

1) Tools under consideration

* Smithsonian Transcription Tool
* FromThePage
* Transcribe Bentham
* ALA Biodiversity Volunteer Portal
* Transcribr (National Archives)
* T-PEN

2) General questions for discussion

* How do we load images and output completed transcriptions
* How will we deal with multiple transcriptions of a single page? None of the tools are designed for this.
* Other issues?

Ruled out

1. Transcribe Bentham – learned at Webwise that code was nightmare and difficult to get data out, our own experience was navigation wasn’t great (advantage – TEI markup toolbar but we may not make use of it)
2. TPEN – least user friendly, steep learning curve so not easy to jump into

Narrowed down to 4 options

* Smithsonian Transcription Tool – like tool, user friendly, polished, Big problem is that code is not available and tied into SI systems. Can’t do a install in time for project. We asked SI if we could provide them content and they host. Carolyn is investigating. Process - We would send JP2s and upload to their DAMS and they would give us output. Our concerns are with beauracracy of SI and time to implement. Might be better as long term solution. Trish asks Mike What are out technical needs if we go this route? Mike says there doesn’t seem to be any integration points, and code being closed is difficult to create any. We’d want to replace our copy of OCR when we get it back – exports in XML, or txt would be both fine as formats. Wait til Carolyn gets back to us next week before ruling out.

3 Other options are ones would have to host at MOBOT. Need to allow time to work with IT dept here and get it set up.

* FromThePage – written in Ruby on Rails, new to Mike but John and Joel have some familiarity. Ben Brumfield working on TEI/XML exports (making upgrades spring and summer). Negative – development is 1 person, Ben.
* ALA Biodiversity Volunteer Portal -We can submit content and they can host – they can create an expedition for us. Their code is open source so we have option of taking and creating our own. Code in Groovy.
* Transcribr (National Archives) - Tool that Harvard has used. Hard to determine how it works because most of the content is already transcribed. Drupal platform

Martin mentioned community support at SI as being important – Trish asked what does community support mean? How are users engaged and kept come back? There has to be people managing this -It’s the staff around the tool that manage the community.

Allowing for Multiple transcriptions of a page – which tools support this? Ben Brumfield is not aware of any tools with this functionality. He has a blog post about the challenges of this.

Page coordinates – lack of support in all the tools with exception of TPEN which has line level support. William says we Could do after the transcription process. We’ll have to look into further but doesn’t hold us up on our decision.

Next steps:

* Feedback on tools from Joel , Jenna, Francis by meeting next Wednesday
* Mike and John will look at code for FromThePage, ALA, and Transcribr