

This is the shared Google Doc for Kirstyn Leuner's talk  
"Hands On: Producing the Stainforth Library of Women's Writing"  
4/7/15  
Part II

**For Part II, we're going to break into groups of 3-4.** For each question, group will discuss and record their conversation for 5-10 minutes on each topic listed below. We'll go around and each small group will share some approaches with the larger group. Then I will share mine before we move on to the next question. (There are 4 questions or challenges like this.)

**Challenge 1. You have a 600 pg 19<sup>th</sup> c. library catalog ms to transcribe. [Here it is.](#) How would you handle the transcription process?**

- In what platform or using what software?
- What would your editorial philosophy be? In other words: how faithfully would you want to transcribe every mark on the page?

Would you encode it during transcription or let the transcription stand alone?

[Type your group responses, ideas, questions here]

[Laura and Ridie]

platform: <http://www.juxtaoftware.org/>

yes, transcribe every mark and encode as we go!

- transcriptions for Occom: Using Juxta (learned at ACRL) - can have image of the scan in one window, and the other window for transcription;

What are the goals and audiences for the representation of the data?

How do you anticipate how people are going to use it?

At the beginning: how do you know how much time and effort you want to put into this?

Let people's use cases drive the direction of the markup.

Do you want every stain inkblot and tear?

Do you need the transcription at all? What you really need is a database.

Database - use the database as the front end, and point to the manuscript for location of the mention of the name.

Shelfmarks aren't true shelfmarks

What do the shelfmarks tell us?

Shelfmarks might be a classification system  
Focus on the content, not every wrinkle and tear.

The image facs is so good!

Maybe J10 is a case?

Was he shelving as he acquired it?

Jay notes that he doesn't mark acq dates or enough format marks 4o 8o

section of site: projects made with this  
feedback  
crowdsourcing  
doi

-----

**Challenge 2.** You're excited. You want to publish the data before you build the interface and link the raw transcription to any further resources outside it (like personographies found elsewhere on the web) or even before you cross-link related titles and authors within the archive. How would you do this?

- How would you edit the data? You have about  $500 \times 24 = 12,000$  lines of data.
- Do you edit every transcribed line or do a random sample edit?
- Do you do some encoding when you edit, or just leave the transcription alone?

- Where would you publish the data? Github? Host on Omeka? Do it on the institutional server with an easy interface, such as sites.dartmouth.edu?
- What would you want those files/that file to look like? What file types?
- What questions do you want to be able to answer with this data?

Helpful Links - examples to help you brainstorm solutions or questions:

Page 18 (interline ex page) in Ms: [PDF image](#)

Page 18: Transcription in [.xlsx spreadsheet](#) in Google Sheets

Page 107: in Ms: [PDF image](#)

Page 107: Transcription in [.xlsx spreadsheet](#) in Google Sheets

page 108: in Ms: [PDF image](#)

Page 108: Transcription in [.xlsx spreadsheet](#) in Google Sheets

[type your group responses/ideas/questions here]

### ----- **Challenge 3**

Scenario 1: The first scenario is based on reality and pragmatism. For the first scenario, I would like you to gather together links of digital archives that you know -- either you've helped build them or you know of them -- that might be a great model for what the Stainforth digital archive will look like and work like in the future. Bear in mind that there will be approximately 2 full time people doing this. Me and one other person. Whoever that is.

Scenario 2: For the second scenario, go big and dreamy. I want you to think about what those digital archive cannot do that you want the Stainforth library to be able to do. What functionality would be the most useful or cool -- nevermind that it might take an astronaut -- or similarly talented person -- to program that functionality.

For both scenarios, assume that we have finished collecting authorities data for all of the authors and titles in the library. Every title is linked to a digital edition, and every author has an authority record where there is one. Those authors that are unknown are tagged as such. Sketch or brainstorm for 10 minutes as a group and then we'll discuss what you've talked about.

[Type your group comments/ideas/questions here]

---

#### Challenge 4

4. We have to maintain this project for some time. How would you do this?

- How do you keep the data and the site on the radar for teachers/students/researchers?
- How often do you perform “maintenance” work and what is that to you?
- How do you plan ahead for the website’s demise? What do you think the shelf-life for the project will be?

[Type your group comments/ideas/questions here]

Mercantile Library of Philadelphia

Accession record: when they bought it and who from  
when they de-accessioned it and why and to  
dealer from Color

*Thank you for participating and for sharing your ideas!*

*-Kirstyn Leuner*