**IMLS PURPOSEFUL GAMING MEETING**

*February 11, 2014 (Baltimore, MD)*

**ROLES**

*Joe* - MCZ work on transcription. William Brewster (4000 pages on BHL for trans). Primary role: identification and review of various transcription tools and guiding selection of one for the project.

*Trish* - Principal investigator: responsible for keeping track of who's doing what, primary contact with IMLS, interim reports, main contact

*John* - Minor role. Provide digitized nursery and seed catalogs. Paul Silverman has come back and is working 3 days a week part time. Previous grant from NEH to catalog material. John's role is to supervise.

**NOTE**: TAG Group - Weekly call for transcribing so a tool can be selected

*Marty* -  Fairly small role - providing digital content from seed nursery catalog from Cornell. Primary person involved. Kevin Nixon (taxonomist) is also on grant. Others will be involved in digitization. Current plan is to outsource the scanning to Trigonix (Montreal). Structuring - Pagination, markup, appropriate tagging. Hesitant to digitize serials - BHL can't handle article-level metadata.

*Mike* - Scanning seed lists and generating sets of OCR. Probably not much need to coordinate with Cornell and NYBG: seed lists vs. seed catalogs

**PROGRESS**

Cornell

Copyright issues: International copyright cutoff dates - set by institution doing the digitization. Pre-1923 vs. 70 years after author's death

Got some publicity. Last week, library staff members gave presentation at Horticulture meeting - generated some interest. Spring semester: Dan Steward (IMLS intern) working on things with Marty, coordinating with NYBG and working on selection process for the most appropriate materials. Start out by identify catalogs for grapes & list of firms. Also interested in "kitchen garden". Coordinating NYBG & Cornell meetings.

National Agricultural Library has been digitizing catalogs and putting into IA and now into BHL. Marty set up collection in BHL for these types of catalogs. Now 2000 items in this collection. But not all that are in IA and Mike is working on importing those. Marty talking to NAL about coordinating what they are scanning. They are targeting large firms and women-owned. This won’t affect how much Cornell digitizes for PG project but just what they digitize. Tom Blake of Boston Public Library approached him and asked if BHL wanted more nursery catalogs. He knows of some at MA Horticultural Society. He gets state money to digitize collections around the state. Challenge of comparing and deduping records is still a problem that BHL hasn’t solved. Collection is cataloged in a variety of ways. Larger firms have individual catalog records but holdings are listed in detail. Firms with fewer catalogs may not have its own catalog record for the firm. They are put in Misc group and physically grouped together by alpha name e.g. Misc A-D. Have to rely on catalog dept to agree to create those. Marty is now working pt not ft and will be for next 3 yrs until he retires. Tuesday, Wed, Thursday are his working days.

MOBOT

*Trish, William, Mike L.* - Service agreement work, press release stuff at MOBOT (front page article for STL Post-Dispatch), writing RFP draft, looking at game design (what are the essential elements needed), general project management

*Mike* - Waiting for cataloging records to be created so scanning can begin. (Check with Doug and Jason on selection criteria)

NYBG

Coordinating what to scan. Setting up mechanisms to try to not duplicate each other's work as much as possible. Mellon money for two scanning technicians. Leaf Aptus camera mounted on Tarsha copy stand. Might have already started scanning.

**RFP**

*Marty*

* Suggestions: Game app should be a web vs mobile (more universal)
* If it's successful, continue to use it as more content is added
* Do what we can to reach the widest audience

*Trish*

* Having something open source would be ideal – Mary Flanagan of Tiltfacter Lavbs recommended HTML5 (can be repurposed for mobile)
* What can we expect once it's deployed?
* Maybe people could take it and continue to develop it further

*John*

* People expect a purpose-built app that can be downloaded from app store.
* What is the life expectancy of the product? How long do we intend to keep it going?
* HTML easier to support
* Do we get the source? Even with the souce for mobile apps, we would need to develop a level of expertise and distribution. iOS and Android will require a lot more resources
* IOS platform may limit where we can provide access – Apple limits the distribution model to their app store.

\*\*Remember goal of the project is to test the effectiveness of gaming for text correction and NOT to produce a game that others can download and reuse in their libraries.

Need to maximize the audience.

Have game ready to go live by May 2015, run for a set period of time to evaluate the effectiveness

**NOTE**: Need more time to discuss the topic. Continue conversation in next phone call.

**TRANSCRIPTION TOOLS**

*Joe* - Up to this point we've narrowed the tools down to about 5 or 6.

*Requirements*

* Open source
* Crowdsourcing
* User-friendly
* Admin oversight
* Admin editing
* Page coordinate data
* Sustainability
* XML Exporting

T-Pen

upload limits, not as user friendly (somewhat of a learning curve)

FromThePage

*Advantages:*

* Import from IA (works well)
* Semantic mark-up for subject indexing (can easily creat & manage ontologies). Can be done later by editors.
* Transcriber notes/comments
* TEI supported for transcription exports
* Good support (Ben Brumfield)

*Limitations:*

* Hosting charge
* Sustainability
* No page coordinate data

*Plans (this spring and summer)*:

* Upgrading Rails framework
* Upgrading login and commenting systems
* Adding type-specific attributes to subjects (lat/long for place names, etc.)
* Complete Omeka integration and TEI-XML export
* Switching encodings to Unicode

Ben is available to do custom development and installation on a contract basis, though limited through summer 2014.

Transcribe Bentham

Ruling out due to code being a mess (quoted today in a talk on transcription)

ALA Biodiversity Volunteer Portal - simple, free-text entry. can rotate images in viewer (good for marginalia). Transcriber notes/comments. Not sure if we can upload enough content for trial purposes. (We would have to host it ourselves.) Designed for biodiversity sources. Special fields for species occurrence data

Transcribr (NARA) - Drupal, Advantages: difficulty rating (beginner, interm, adv). Browse by difficulty, year, and transcription status, simple free-text, transcriber notes/comments

Limitations: no page coordinate data, very little on-transcribed content for trial purposes

Smithsonian Transcription Tool

Advantages: simple, free-text entry, ease of use, good navigation (browse by A-Z, Popular, Latest, Completeness), Registered users can transcribe & review (project staff make final approval)

Limitation:  Code not yet released for general distribution

*-* Trish: Martin would like it to be our lon

- Joe: How hard would it be to pull material from BHL into this transcription tool?

- Trish: Rule it out?

- John: Reject it for the short term, but maybe accept it long term? (going through Smithsonian can be a bureaucratic process) Better to have something we can install ourselves rather than depend on SI.

*Trish* - What are the pros and cons of hosting the tool ourselves or through someone else? If we host it, it will be at MOBOT. T-Pen (SLU) would charge to have it developed further. Advantage is that it's line-level. Don't need markup for the purposes of the game. Short timeline for getting this tool in place.

*John* - This tool needs to be integrated with BHL.

*Joe* - Opinion: FromThePage or ALA would be best tool

*Trish* - ALA's interface is strange with transcription box below the image – something to consider