**Purposeful Gaming planning meeting 11/20/13**

Present: Susan Fraser, Joe DeVeer, Mary Ochs, Marty Schlabach, Willliam Ulate, Mike Lichtenberg, Trish Rose-Sandler, Doug Holland

Meeting minutes

1. Review narrative – does everyone understand the goals and objectives?  Are there questions?

Q: Who will review bids that come in from gaming company and help in selection? Good to include 15-20 yr age range on review team, different age levels should be represented. Potential candidates include students from STL universities, students from partner universities, our kids?

1. Review workflow diagram – any questions?

William explained the details of the diagram. It helps us develop automated ways to create groups of content and then filter them out by quality. We want to identify ones that aren’t working for OCR then push those to transcription. The number of non-alpha characters will help us determine this. We need at least 2 outputs for each page so we can do automated text comparison. The comparison will generate acceptable accuracy scores. Those with alot of errors will get manually reviewed. Those with acceptable score we will push to the game. The game will have users review content at the word level. Several people will review the same word so we have confidence in their feedback. We will then assemble the corrected text output. Typical errors we see repeated in other pages. If we have a valid statistical measure we could apply this to other pages in corpus. Any page we review manually we will add to a ground truth set of files. We can use these files to measure against and obtain accuracy scores. Using “post processing error correction system”. It is Semi automatic – system proposes errors to humans. We may discard some outputs as being too problematic. Language can introduce challenges.

Q: Different OCR software has different strengths – what will we be using? Tesseract maybe, Prime OCR. Are we comfortable with accuracy of those software? Mike L says it does better than OCR software used by IA because IA optimizes for efficiency and not accuracy. Also IA was not using latest version of ABBYY software last we checked.

Q: Will MOBOT staff be responsible for doing the manual reviews at MOBOT? Yes, we have Some staff dedicated to this but could use more help with this. Some partners said they may be able to help with this when they have downtime from other project responsibilities.

1. Review schedule of completion – does everyone understand and agree to the timeline?  Are there any changes that need to be made?

Yes understood and agreed. No changes to make

1. Review budget – can everyone still meet their staffing obligations and cost share?  Are there any staffing changes?

Staffing changes reflected below

1. Administrative details – has everyone notified their controller’s office of the grant?  Is there any more paperwork we  need to do there ?

Susan Fraser said her office needs contract/purchase order or service agreement from Missouri Botanical. Does anyone else need this? Cornell will need a subagreement/subcontract.

**Other topics**

Frequency of meetings during project? Marty doesn’t think the larger group needs to meet more than once a month at beginning. But there will need to be some smaller groups meeting more regularly, particularly as they need to coordinate what to scan and not have overlap.

William said the Unconference software we used today is typciallly used by MOBOT for international calls so we will use another one for national calls going forward.

Q: Can partners use a MOBOT number for smaller group calls? Lets see how the needs play out. It would be preferable to use skype or other free conferencing for small group calls and communicate via phone and email. Cornell has conferencing calling capabilities . If other partners have conf call needs that can’t be met by their institution we’ll investigate some options.

Q: Does everyone understand and agree to upload to IA and make available in BHL? Yes,

Hort catalogs – Marty would like to set up a subcollection in BHL.

It’s important to keep other BHL staff, outside the project, updated on our work in case it will impact their workflow

Marty asks could we developing naming and verification at the cultivar level? He isn’t aware of a universal cultivar list we could work with.

Attending Webwise meetings – do we need someone from each partner to attend? Mary felt it was not particularly useful for academic libraries, its mostly a platform for IMLS recipients to talk about their projects and network. Not every project does a presentation, there are some poster sessions, IMLS determines which projects to present. If not asked to present we should request it. \*Advantage: We could hold our in-person project meetings there if we all attend.

Date for Kickoff meeting? December or Jan? Better to do early December so we know what we should be working on.

**Staffing/Tasks**

**Cornell –**

2 Staff - Marty and Kevin (only people named in grant) Marty facilitating overall product, Kevin would like copies of output for his own textmining work. Other staff not named on project - Carol Lowe – she does structuring and markup like Macaw work, she will be involved in beginning up to structuring process. Francis Web was initially designated to help with IT but her job is changing so may have someone else do this. Also a conservation person involved to determine suitability for scanning.

Primary Tasks:

1. choose hort catalogs
2. scanning horticultural catalogs
3. uploading to IA for ingest by BHL

Cornell will coordinate with NYBG on selection

Clarify total to be scanned – says 70k in specifications digital document. Budget justification doc says 370 catalogs with 125k pages . \*Marty will look into. All money is towards digitization so price based on quote from company.

Q: They did not include travel for Webwise in their budget. Will they be attending? Somebody from Cornell can go if we need them to.

**NYBG**

|  |
| --- |
| 6 staff - S. Fraser,Director-Admin,  |
| J. Mignault, Systems Librarian |
| K. Nolan, Digital Project Manager (won’t be on as cost share but will utilize someone else) |
| L. Studier, Metadata Cataloger |
| Y, Choi, Catalog Librarian |
| A. Tschinkel, Scanning Tech (he’s on another grant so will hire someone soon)Tasks:1. Selection of hort catalogs
2. Scanning horticultural catalogs

15k pages? Not sure how came up with number but based on a parttime person for 1 year.**Harvard**5+ staffConnie RinaldoJoe DeVeerRobert YoungSpecial Collections Asst. Interns (probably 2 will be hired in Dec)Tasks:1. Selecting transcription tool
2. Implementing transcription tool
3. Promoting tool to public
4. Hiring Interns
5. Using public and interns to transcribe field notebooks and hort catalogs
6. Using Mechanical Turk to transcribe field notebooks and hort catalogs

**MOBOT**7 Staff  |
|  |
| TrishWilliamChuck MillerMike LichtenbergDoug HollandMike BlombergInformatics Architect |

Tasks (Library):

1. Select seed lists
2. Scanning seed lists (8k pages?)

May need to do more due diligence on copyright

Tasks (CBI)

1. Hiring gaming company
2. Working with them to develop game
3. Building tech framework for comparing text output
4. Assigning accuracy scores for pushing to gaming tool
5. Deploying game
6. Evaluating accuracy scores against ground truth pages
7. Generating error matrix
8. Writing reports, dessminating findings