

IMLS Civil War Newspaper Project Conference Call

July 14th, 2004, 9:30 AM

1. Update from IMLS Newspaper meeting at ALA

1. Rachel Frick of Richmond, Kenning Arlitsch from Utah and Brenda Bailey Hainer, CO, met at the Newspaper Digitization Interest Group meeting at the ALA annual conference in Orlando. Overall the interest group was a great way to hear about other non-IMLS sponsored projects. Information about the newspaper digitization listserv and Richard Urban's website <http://www.cdphheritage.org/resource/newspapers/>. Also announced was the upcoming NEH funding opportunity. Dovetailing with the completion of the NEH funded newspaper microfilming project, the NEH announced that it would launch a new project aimed at digitizing "modern 20th century" newspapers. They are seeking 10 partners to work in a 2 year pilot project focusing on newspapers from 1900-1910. A formal announcement and project specifications are to be posted on the NEH website "in July". It was discussed that the data for this project would consist of straight page images and only page level metadata. This was met with much consternation from the group; of the 16 active projects represented – all were providing article level metadata. After the general meeting, the IMLS project representatives met and discussed the status of our projects. Both Utah and Colorado stated that their projects have moved from research to production, and they had lots of content to work with. Colorado is working with OCLC to conduct newspaper digitization workshops/ seminars, with the idea that this would be a traveling educational opportunity. The first workshop would be in Colorado in the late fall, and second in the Northwest U.S., mid-winter, and on the East Coast in the spring. Details are tentative at this point, and will firm up after the NEH official announcement. OCLC expressed an interest to work with Richmond in conducting the East Coast workshop, and it was suggested that it would be beneficial to possibly involve ACS or Solient. Kenning and John Herbert authored an article entitled "[Microfilm, Paper and OCR: issues in newspaper digitization](#)" published in *Microfilm and Imaging Review* v.33:2. Greg Crane speculated that the representative from the NEH was Helen Aguara (sp?) accompanied by a young man named "Fields." Crane introduced a new staff member, Alison Jones. The July Conference call was cancelled due to vacations; we will reconvene via telephone in September.

2. Status of content and tagging production from DDD

Tagging is better, although it is still evident that they do not understand TEI – Spelling errors are a residual problem that should be corrected as more content is processed. Also Jason of DDD has made some changes in the Q & A process to address the higher than standard error rate. It should take DDD 10 months to completely process the Dispatch.

Crane clarified that the OCR to be done by Tufts to fulfill the cost-benefit analysis portion of the grant (for which Tufts is responsible) needed to be done from originals, and would be done from copies of *The Christian Register* which they have access and rights to.

3. Discuss current state of project, adjustment to timeline, and clarification of responsibilities
It was agreed that an adjusted timeline and responsibilities document should be created and distributed through the group – first draft by the 1st of August – edits and corrections to be made by mid month in order to give to Jim Retting, PI at Richmond, for incorporation into the 1st annual report.
4. Metadata meeting in Boston
Andrew will work with Greg directly to coordinate this meeting. At this point in time it is not necessary for anyone else from Richmond to attend – the target date for this meeting to occur will be the last week of July.
5. Fall Seminar
Bob is working on invitations and agenda – has been on vacation – will have a clearer picture by the end of the month – wants to meet with others in the U of R group to gather what data will be available to look at, and other topics for discussion at the seminar.
6. Other business
FEDORA web cast is set for Friday July 16th at 10am EST. This presentation will be archived for one week. Members of VIVA, RALC, Colleagues from Johns Hopkins University, ACS students, and members of Pat Schoknect's Frye Leadership Institute Class were invited to attend in person or virtually.
Pat Schoknect asked if UR or Tufts would be willing to host the eastern workshop of the OCLC mentioned above (in 1); there seemed to be willingness, depending on sources of funding, for either UR or Tufts to host, or possibly to host it in DC.
Greg mentioned utilizing version tracking software (CVS) to manage and document changes in our content files. This suggestion was wholeheartedly supported by the group. Tufts will manage this aspect and work with U of R on the implementation – this will be addressed when Andrew goes to Boston.