

IMLS Grant Conference Call—November 9, 2005

1. Contingencies for delivery of Richmond Daily Dispatch to Perseus: This is primarily contingent on when we get our last shipments of the files that have been corrected by DDD. We have finished CHECKING all files from DDD, and I will be sending the last shipment of files to be corrected by DDD later this week. DDD are still working on corrections they have already received from us. But it should be a matter of weeks rather than months when all files sent back to DDD will have been returned to us. Once we have those files back from DDD, the files returned from DDD have to be reconciled with the version of the file as it existed in the main SVN repository (and as it still exists). Our students have already started on that new workflow. Because this workflow relies on the files being compared (they're using "Find Difference" command, like the diff command in UNIX, to compare files), being VERY similar, this workflow will become impossible if either files are substantially changed before the DDD changes are added into the files in the "canonical" directory, we have to add those changes from the corrected DDD files before apply global changes to them. That said, assuming DDD can complete the remainder of their corrections relatively quickly, we will be able to add in those changes on our end quickly as well, making it as matter of weeks, not months, before the Perseus transformations can be run on these files. Andrew Rouner explained that we have a Subversion repository for sharing XML files between UR and Tufts. Greg Crane has a temporary user name and password so he can access it from outside the UR network.
2. Rachel Frick and Andrew Rouner Rachel and Andrew have been invited to turn their Educause presentation into a paper, and they have accepted, They will roll in the material Alison sent prior to Educause and credit her on the paper.
3. Nancy Woodall reported that she has convened meeting to bring together pertinent UR staff whose technical expertise is needed to move the project forward. Among the things they bring to the project are expertise in Web interface design and activating on our Web site a survey modeled on the Colorado and Utah surveys. The most significant issue is to use MySQL rather than Oracle with Fedora. UR needs to move project files from a server that is at 95% of capacity. There is a plan to do this.
4. UR has done everything we are supposed to do to make DXLS work. We have found that the Perl module it was calling on is from a different release. The problem is in the software, not the local installation. The developers will have to fix this problem. We hope it will be solved in the next few days. We hope to have the Richmond Dispatch files and survey up by the end of the month.
5. Bob Kenzer and Jim Gwin have met twice to develop the Web site's contextual information. Their goal is to get this to Andrew by Thanksgiving. There will be similarities to the Brooklyn Eagle site. Bob is developing a time line that will show both local Richmond and Virginia events and national events.

Bob raised concerns about the potentially “schizophrenic” nature of the Web site if we call it broadly a Civil War newspapers site. The overwhelming preponderance of content will be Richmond content’ that includes only one Civil War newspaper. Careful design of the site’s front page can offer appropriate avenues to various audiences can help solve this problem. These audiences include historians, Civil War buffs, genealogists, the digital library community, K-12 teachers and students, and more.

6. Greg Crane reported that he has reached a point where he can focus on indexing issues in the Dispatch. Newspaper articles do not offer the continuity in topic that monographs usually offer. This necessitates changes in approach. His files not hold:

- 23 million words
- 6.7 million annotations
- 520 million annotations
- 520,000 surnames
- 420,000 place names
- 61,000 dates

He is trying to allow people both to search and to browse people, places, and dates. Response time is an issue, as is the user interface.