

IMLS Grant Conference Call—March 10, 2004

1. Greg Crane reported that he had not been able to open the Byte Managers images Rachel forwarded to him. Rachel said she would make them accessible through a Web page.
2. Rachel Frick reported on the cost figures quoted by Byte Managers. To create images of all pages in all three papers, do OCR, clean up the OCR, and do TEI lite tagging would exceed our budget of approximately \$130,000 for these processes.

Greg Crane reported that Project Perseus enforces a lower accuracy rate for OCR than for double keystroking. We might use imperfect OCR output for searching; then we could pick a subset that we raise to a higher level. The question is: How do we select that subset? Potential answers:

- Keystroke only news stories and do other parts in OCR only.
- Treat ads selectively, perhaps culling them from the *Richmond Dispatch* only once from each week's issues.
- Supplemental materials identified to enrich the Richmond paper are held in the UR collections (Rights issues would have to be negotiated for some; others are in the public domain. See table at <http://www.unc.edu/%7Eunc/ing/public-d.htm> to determine which are in public domain.)

Given problems with the quality of the microfilm version of the papers, especially with character recognition in each page's column closest to the binding gutter, Jim Rettig raised the question of whether or not it would be preferable to scan the originals on the *Richmond Dispatch* in UR's collection. A planetary scanner might not be able to capture the entire page since these papers have a very large trim size. We should ask UVa what it could do and at what cost. Some selective scanning from the originals might have value in comparing cost and quality.

Bob Kenzer said that if we can afford to digitize only one newspaper to the highest level, it should be the Richmond paper. He can most effectively organize a conference focused on Richmond; the existence of the Accessible Archives digital version of the Richmond *Inquirer* makes it unlikely that anyone will digitize the *Dispatch*. [An examination of the budget shows that there are no funds earmarked for this conference!]

Greg Crane suggested that there may be a Boston weekly that would fit within our budget and that could substitute for the Philadelphia paper. Bob Kenzer suggested that if we make a switch a Washington, DC, paper might be a good choice; that would give the views of the war from both the Union and Confederate capitals.

3. Supplemental materials—Bob Kenzer and Him Gwin distributed copies of two documents, one listing supplemental sources related to Richmond, the other listing supplemental sources related to Philadelphia.

Bob Kenzer raised the question of incorporating information at high quality, relevant Web sites not affiliated with an institution that can be presumed to be a stable, trusted

repository. Perhaps we can offer them such a home through our repository strategy. Among such sites are some that digitize a single family diary from the era.

The project offers opportunities for workshops for local historical organizations. These can, for example, help those organizations develop strategies to make their materials more widely available and more useful to diverse audiences.

Greg Crane cited the value of applying GIS to maps to give them added value in combination with information in the newspapers and the supplemental materials.

4. Jim Rettig reported that his need to devote considerable time to a family matter has prevented follow-up on the assessment workshop he and Rachel Frick attended in early February. They did receive the promised feedback from the workshop facilitators; its was less informative than expected.
5. A group from UR will travel to Johns Hopkins later this month to discuss repository strategies.