

Approved, December 14, 2006

**Digital Advisory Committee (DAC)
Meeting minutes for 10/26/06
Recorded by Helene Tieger, Bard College**

Present: Jennifer Palmentiero, Tessa Killian, Zack Spalding, Chris Pryslopski, Janet Dymond, Jane Ward, Laurie Shedrick, Bob Clark, Helen McLallen, Sally Pellegrini, John Ansley, Duane Watson, Helene Tieger, Judy Fischetti.

Introductions.

DACs chair Helen McLallen began the meeting at 9:35 a.m. by welcoming new members. All introduced themselves.

Purpose and Role of the Committee.

In the absence of John Shaloiko, Tessa discussed the purpose of the day's meeting: to review activities of the DAC, and to provide future direction for the Hudson River Valley Heritage (HRVH) service. She provided a brief history of the committee and referred those present to two documents available in the "About" section of the HRVH site for more extensive information. Distinguishing a 'working group' (a group, not necessarily comprised of DAC members, who meets two or three times to achieve a desired outcome) from a "subcommittee" (a long standing group dedicated to working on issues over time), Tessa described the need for establishing the former to work with SENYLRC on indexing issues and revising the scanning guidelines.

Update on Recent Activities

- Library Services and Technology Act (LSTA) grant – two year project, to be completed by March of 2007. Intended to provide training to regional libraries and historical societies for creating individual digital projects. The result would be—and is-- a digital repository of images related to the Hudson Valley, and hosted by SENYLRC. To date, the scheduled training workshops have been given, and 19 participating organizations have contributed content. The last phase of this project will be to provide members with packets of promotional materials, and to document the promotion of all of the collections. Bob Clark recommended utilizing specialty listservs to promote the collection, and Chris Pryslopski suggested posting related press articles on the HRVH website.
- New training program - Tessa and Jen have condensed the content of the LSTA workshops already given into a 2.5 day program, to be offered at the end of November. The cost will be the price of membership in SENYLRC. A full day will be devoted to issues of copyright and management of a digital project; another will cover metadata and use of Content DM. The projected plan is to have them provide this training once or twice a year, as needed.

- Usage Statistics – Zack presented a brief summary of HRVH use. Since May of 2006, there have been a total of 65,000 hits from 40 countries. A spike in June (almost 23,000 hits) reflects the Annual Meeting with surrounding press coverage. One quarter of the hits originate from Google searches, while the use of ‘RSS feeds’ and Wikipedia links for HRVH embedded in the Wikipedia entries for Bard, the Huegenot Society, and the City of Newburgh also account for a good number of hits. Currently Zack is working to enable individual member institutions to search their own usage statistics.

Digital Newspapers

Tessa provided the background on efforts to explore HRVH’s role in providing digital access to regional newspapers of historic significance. The enormous popular success of several North Country newspapers (over 1 million hits per year over several years) convinced SENYLRC that newspapers should be provided as part of their services using HRVH as the platform. SENYLRC staff Jen and Zach worked with an outside vendor, using Kinderhook Rough Notes as a test title. Results were not good: OCR quality was very poor, with almost no legible words resulting. Hudson Micro Imaging was then enlisted, and results were much better. This vendor used 24 bit color, not grey scale, in the initial scanning process.

Discussion ensued about process and pricing: scanning from originals vs. scanning from microfilm, with costs ranging from \$.59/per page, to \$1.00 per page (microfilm created from originals, scans created from microfilm), to \$1.54 per page (scans made from original paper, preservation microfilm created from digital objects). Sample images were viewed. SENYLRC leaning toward second option. Content DM also poses additional OCR problems, as it only reads the last column of text on a page. Given this as well potential licensing problems related providing access to newspapers through Content DM, John Sheloiko is thus considering another platform as the provider of newspaper images on HRVH, though other images would remain on Content DM.

Discussion also followed re: the necessity of the OCR function. Bob reported that OCR is not much by patrons at the FDR library. In addition, the North Country newspapers have very poor OCR quality, and yet the popularity of the digital newspapers has remained high over a several year period. The consensus of the group present seemed to be, however, that since indexing and the resulting word searches originate from OCR text, this function is essential to the HRVH project. It was further agreed that quality should trump quantity: fewer items scanned with greater usability would be more useful for our purposes than scanning thousands of documents with poor OCR retrieval.

National Digital Newspaper Project. Tessa reported that the National Digital Newspaper Project grant (coordinated jointly by the Library of Congress and the National Endowment for the Humanities) applications are due Nov. 1st. New York State’s grantee will be New York Public Library. SENYLRC had hoped that the NY State library would have applied for this, and that given the background work with HRVH, that SENYLRC

might have worked with the State Library-- but only one organization per state is chosen. Tessa is exploring other grant options to pursue the digital newspaper project.

New LSTA Grant Application

Tessa reported further that SENYLRC has been selected to apply for a \$17,500 invitational LSTA (Library Services and Technology Act) grant. Applications for this grant are due January 1st, 2007. In an effort to broaden HRVH use in the region and beyond, Tessa will be working with teachers and teachers' centers to ensure that this grant application will cover state curriculum requirements for multiple age levels. Grant money would be used in part to hire a consultant to write HRVH curriculum with the result that HRVH would become a significant research tool for teachers, students, and home-schooling families.

Indexes/ avenues for searching HRVH (themes, subjects, topics)

The need to increase searching options to HRVH collections was discussed. Citing preliminary documents, Tessa explained that originally HRVH collections were conceived to be : a.) historical; b.) of enduring research value; and c.) broad in focus, and the guiding principle was the Field of Dreams notion : "If we build it, they will come." With a base collection, and initial support from the general community, however, HRVH is ready to move into a new phase of refinement for both access to and content of this collection. New avenues for searching would help to focus attention on the kinds of collections desired, but not yet included in HRVH, and would also help identify untapped user groups. Thus, key questions are : 'What kinds of information is needed?' 'Who will use these collections?' and 'How will people search to find the information that they want?'

Jen indicated that individual institutions would always be available from the home page so that patrons of these organizations would always have the option to refine searches to their home library or historical society. Thematic searches, however, would enable a user to draw from all collections for information on a given topic. While the search box currently available does search across collections, it is desired that topics be suggested to the user to better facilitate research and results.

To begin this process, it was suggested that subject headings (Library of Congress [LC], Thesaurus of Graphical Materials [TGM], as well as local headings) be exported to identify common searches.

A working group was established to further discuss this problem. The group was tasked with suggesting possible search categories, identifying potential materials for inclusion in HRVH, and also identifying possible local subject headings that better describe the HRVH collection than the broad LC or TGM subject headings. Volunteers in this effort included: Duane Watson (Wilderstein), Bob Clark (FDR Museum and Library), Helen McLallen (Columbia Historical Society), John Ansley (Marist College), Jane Ward

(Woodstock Public Library), Janet Dymond (Daniel Pierce Library), and Sally Pellegrini (New City Library).

Scanning Guidelines

Jen described the need for formal guidelines for use of the JPEG2000 file format. A working group to provide these was suggested, to include Jerry Henry of Newburgh Free Library, and Deb Begley, recommended by Laurie Shedrick of the Mid-Hudson Library System.

Wrap-up

Next meeting was tentatively set for Thursday, Dec. 14th from 10-12. The meeting concluded at 11:50.