

DISC Minutes May 7, 2008

Agenda

Present: Beth Black, Wes Boomgaarden, Tom Cetwinski, Tschera Connell, Amy McCrory, Sarah Murphy, Dan Noonan, Sally Rogers, Eric Schnell.

Guest: Melanie Schlosser

Minute taker: Tschera Connell

Implications of recent conference information for DISC activities

Digital Library Federation Spring Forum (Melanie Schlosser)

Melanie identified three themes from the conference and gave examples:

1) **Providing access to library collections and services in small and limited formats (iPhones, SMS, etc.).** *NC State U* library services available in a text-based, small-screen-friendly format: Catalog; Page listing currently available computers; Today and tomorrow's opening hours; Campus directory; Contact information for key library services; Links to external service providers; Status info for U buses. *DataDyne* is exploiting SMS-based mobile phone capabilities to improve public health services in developing countries. CME, statistical surveys, vaccine reminders, etc.

2) **"Cultural Commons"** – using digital objects to create culture. *Smithsonian Institution* is developing an interesting "web strategy". It's based on creating a "Smithsonian Commons" which is a big pool of digitized content surrounded by a "halo" of services. They are decentralized, like we are, and knew that a top-down solution wouldn't work. Want to encourage innovation, so voluntary participation. Strategies: Encourage service development by using a software development kit and APIs. Encourage users to reuse content by placing less restrictive copyright statements on public domain collections. Success would be measured by the 'fecundity of user communities.' (*LoC Flickr Commons* project is an example of this. Making collections readily available for use.) *Stanford University* is creating a text corpus of the earliest printed books in English, generated by reusing freely available content from a number of sources. Corpora are usually made up of painstakingly keyed and marked-up text. This includes OCR'd text and lots of other stuff, and will enable a 'distant reading.' Could something like that be done with our Google Books? The unifying idea is that our digitized collections are the building blocks of culture, and we should be focused on making them available for reuse, mashups, etc. Partly a technological thing, but partly a mindset. We are protective of our collections, which makes sense in a physical environment, but not the digital one. We are focused on protecting them from commercial exploitation, but that's not all that's going on out there.

3) **Scholarly Communications.** *UIUC* has accepted a proposal for a scholarly communications department: Copyright services, connect scholars with subject specialists, maintain an IR and offer 3rd party deposit services in other repositories (like PubMed Central), data curation for 'small science', resources for alternative publishing.

Also offering *BibApp*: "Institutional bibliography." An expertise system like Pro, but offers interesting visualization tools and an interface with Sherpa/Romeo. Faculty-focused, not administration-focused. Also, designed with web apps, it's really user-friendly and modular. *IU ScholarWorks*: Moving from just an IR to a "set of services" including open-access journal publishing. Both libraries developed their scholarly communications departments in part to support their IR. Low adoption. Needed a broader context.

Midwest Archives Conference (Dan)

Most useful to DISC is the Sound Directions Project (Mike Casey, Associate Director for Recording Services, IU) This project uses defacto standards for audio presentation. Dan will present more on this topic at the June meeting. Dan also attended a session on shared library/IT positions in the institutional hierarchy. Several small libraries have library directors and IT people who are one of the same. No large institutions have this model.

Coalition for Networked Information Spring Task Force Meeting (Tschera)

The excellent plenary session (Paul Evan Peters Award Lecture) was given by Daniel Atkins, Director of the Office of Cyberinfrastructure, NSF. Dr. Atkins spoke of institutions in transition and called for development of "collaboratories", virtual organizations tapping into diverse talent and resources. Things Dr. Atkins has learned during his career: what is most important is a paradigm shift, not the details, not the brand; job security carries an obligation to take risks and do what needs to be done; have vision and passion (and use administrative power in service); use skunk works to create existence proofs for doubters.

Two other sessions were especially useful. The first, Accessing Research Cyberstructure Needs at University of Minnesota. Participants in the project (members from the College of Liberal Arts (CLA) and IT) held interviews with 4 diverse research groups: 1) biological simulations research group; 2) magnetic resonance imagery; 3) "Saving the world with Bill and Melinda" – this group handles all the worldwide data from foundation harvest project; and 4) a large Social Science Research Center. The commonalities of needs that were identified included storage, accessibility of data for the group, scalability of the groups' activities, and sharability of data. The data suggest a layered approach and supplying interfacing staff that can assist the researchers. The services being designed are a joint venture of CLA/IT. The other session, given by George B. Adams, Associate Director for Computational Nanotechnology, was about the Purdue nanoHub. nanoHub provides users with over 70 simulation tools for research and education accessible from a web browser. In thinking about the work of the Datasets Working Group, Tschera wonders whether this model is workable for other disciplines.

PDF/A progress report (Amy, Beth B., Dan)

Amy, Beth and Dan have been trying to convert vendor created pdf files to PDF/A. Some consistent errors have occurred which they plan to share with vendors in the hopes

of the vendors improving their conformance to the standard. The next step is to recommend directions for DISC. When and if PDF/A adopted, it will be relatively easy to implement for future items. There is a tool, pdfaPilot Server, for batch conversions. This tool was mentioned in answer to question about whether we might do retrospective conversion of files. DISC will need to develop standards for the types of items to be kept in PDF/A and the types of items kept in other formats e.g., tiffs. There will need to be a decision matrix developed. The group will report back to DISC in July.

OhioLINK EAD Finding Aid Creation Tool demo (Amy)

Amy demonstrated the EAD Finding Aid Repository (<http://ead.ohiolink.edu/>). Right now it contains mostly legacy finding aids. A Finding Aid creation tool has been created to standardize finding aids so that a common search will work across all collections. DISC commended Amy for all the work she and others have done over the years. This has been a mammoth contribution to the state, the archival community, and OSU. Amy will send the link to the Finding Aid Creation Tool so that DISC members can experiment as “Guests”. [link: <http://ead.library.kent.edu/>]

Recommendation on how to brand OSU's instance of the OhioLINK DRC (Beth B., Tschera)

Tschera and Beth recommended an interim solution until we (OSUL/DISC) have a better idea how we will use the DRC. The suggestion is to use a branding that imitates the OSUL home page. This will enable us to add the Mills Atlas (if OhioLINK is ready). We can change our branding later.

Any questions / discussion related to items emailed to libdisc since 4/2/08 meeting

- Project updates (Amy sent 4/3)
- Tom's responses to items on Mary Scott's list of projects that were referred to him (Tom sent 4/10).

Some of these projects are in the pipeline; some of the projects have a lot of assumptions, and so Mary needs to talk to faculty; Tom is waiting to hear from Mary.

- New features of DSpace to pursue in near future (Beth B. sent 4/18)
 - New features list sent out
 - RSS feeds are being tested on development
 - Check sums – on production, running at night
 - Currently experiencing memory leaks on production

Project Updates

DARE – zooming not satisfactory in DRC. Raises the question what to do with DARE and MILLS Atlas. More conversations needed with OhioLINK. Nena is pursuing the issues for DARE.

Next meeting: June 4, 2008

Agenda items:

- Audio & Moving Image Standards (Dan)

- Proposed methodology for service gap analysis and creation of joint Assessment / DISC subcommittee (Ruth Gallegos Samuels)
- Products and services list (J:\Committees\DISC\Disc documents\products_services_2008Apr02.doc) (for discussion with Ruth)
- Conference information
 - CIC E-science conference (Dan)

Target dates for discussion items at future meetings:

- UIUC digital curation workshop - July
- delivery and presentation of content
 - adoption of base standards and thesauri - July
- digital project site design - July
- PDF/A update – July
- digital imaging guidelines update - August
- METS – Sept
- OAIS/TDR - Oct