



ART LIBRARIES SOCIETY OF NORTH AMERICA

35th Annual Conference
Sheraton Colony Square, Atlanta, Georgia, April 26-May 1

Session 10: Power to the People: Social Tagging and Controlled Vocabularies

Saturday, April 28, 2007, 11:00 am -12:30 pm

Moderator: Sherman Clarke, Head of Original Cataloging, New York University Libraries

Speakers:

Lauren Cornell, Executive Director, Rhizome at the New Museum

Jenn Riley, Metadata Librarian, Indiana University Digital Library Program

Ross Singer, Library Applications Developer, Georgia Institute of Technology

Recorder: Elizabeth Lilker, New York University Libraries

Sponsor: Duncan Systems Specialists Inc.

Summary:

The moderator, Sherman Clarke, spoke about the focus of the panel, the intersection of controlled vocabularies and social tagging. He said that words go in and out of fashion before we can establish them by our traditional means. He continued, saying that with social computing and blogs, we are all becoming writers and catalogers, and that the basis for both controlled vocabularies and social tagging is literary warrant.

Lauren Cornell, "On Rhizome.org"

Lauren Cornell began by giving an introduction to Rhizome.org, a non-profit organization with a global focus on new media art and technology. Her talk was on one part of the organization's work, Art Base. Art Base was founded in 1999 **[1999 on website, 1996 in my notes]** and is an online archive of new media art with over 2,000 art works. The art works are described by the artists, using a combination of Rhizome's controlled vocabulary and the artist's own terms. There are approximately 180 terms in the controlled vocabulary. It was recognized that these terms alone were insufficient, and in 2006 they began encouraging the artists to describe their own works. She showed several examples from Rhizome where the image was tagged with both Rhizome terms and Artist terms. The terms are

displayed in tag clouds, with the font size indicating the relative popularity of a term. In a few cases, the Artist terms greatly exceeded the number of terms from the Rhizome controlled vocabulary. Cornell said that membership basis of Rhizome fostered the complementary interaction of the two tagging systems and the spirit of Rhizome to allow the artists to describe their own works. She said that Rhizome was considering adding the most popular Artist terms to the controlled Rhizome terminology. She added that scalability was an issue; the process could be difficult to manage at a small organization.

Jenn Riley, "Possibilities for Social Tagging in a VR Collection"

Jenn Riley began by saying that her institution, Indiana University, was in the planning stages for tagging. She outlined the questions they had recently been examining. This included the purpose of tags, who creates the tags, and who are the tags for. She stated that the library world had begun to expand beyond the traditional creators of records, which have included professional librarians, copy catalogers, and vendors. The visual resources community also saw the need to expand. She said that users often know more about a particular resource and that was the metadata and access that was needed.

She talked about the many options of who would be allowed to tag records, possibly including the general public, those who choose to register, anyone with an .edu address, the IU community, specific roles within IU, to specifically authorized users. She also discussed the possibilities of what would be added, and this could include unstructured tagging separate from the cataloger-created metadata, such as Penn Tags, to limiting tagging to specific fields or areas. While libraries are understandably reluctant to cede control of their catalogs, she said there were many options available in allowing users to contribute metadata.

Riley talked about the various incentives users have to do this sort of work, from managing their own personal resources to contribution to the greater good in adding their knowledge to a record. She said that users are already familiar with methods of organizing information, such as Flickr and del.icio.us. Indiana University, like many universities, cannot keep up with the number of images, and creatively allowing users to add content could be a valuable step in expanding access to images.

Ross Singer, "The Communicat"

Ross Singer's presentation on "The Communicat" was on the process he developed for Georgia Tech on expanding collaboration between users, and the library catalog would be one aspect of this collaboration. Singer spoke about the library catalog as an inventory control system, which lacks a relationship to the outside world. Communicat was designed to incorporate social bookmarking and tagging, allowing users to create their own relevant connections to resources and to one another.

Singer said he was not particularly concerned with social tagging of metadata, but the context to which it was added, for example, a physics class project. The project could create relationships between resources such as books, conference papers, presentations, etc. Singer discussed the software that allowed this work, named "GaTher," which saves the items selected by the user and creates associations between them. GaTher uses the Daisy CMS (content management system). Singer talked about the "research trail" that could be created as an "überbibliography." As in Riley's presentation, Singer discussed the various levels of possible input, core: traditional resources selected by the library; community: items added by Georgia Tech groups; and world: items added by people outside of Georgia Tech.

Questions:

Question by Murtha Baca: Where is the user seeing data? Jenn Riley: both from within the catalog and outside. Metadata is contributed from many venues.

Ross Singer: notion of large scale social software is already out there. Systems generally fairly fast, and "quick and dirty."

Question by Eileen Fry: Limited to things we can identify in Ross Singer's talk. With images we cannot do that, no way to share this. Steve.museum [Art Museum Social Tagging Project] comes close to this. the Getty had begun considering the development of a Cultural Objects Name Authority (CONA).

Question for Lauren Cornell: Workflow question on the move of unauthorized tags to a controlled vocabulary. Do they get automated reports at Rhizome? Cornell: haven't done that yet, though they plan to collect the most popular terms, approximately the top 50, then have an open conversation on blog with the artists to discuss which terms are valuable and are most likely to have longevity. Rhizome is uniquely able to have this process. Sherman Clarke asked about the normalization of singular/plural terms. Jenn Riley commented that the problems with singular and plural terminology shouldn't be [put on] the user. The concept is more important than the issue of singular/plural. Machines can do some normalization.

Question by Ross Day on "reverse migration" Is controlled vocabulary relevant? Cornell: Controlled vocabulary will grow, and probably stratify. Rhizome didn't want to lose any words. In the tag clouds, the size of the term represents its popularity.

Question by Vicki Bohm on the research trail. Ross Singer: the research trail in GaTher could be called the "poor person's Web of Science."

Question on incentives for adding tags. Cornell sees this as very important. Another comment that one cannot assume people will participate. Reference to the CLiMB project (Computational Linguistics for Metadata Building) at the University of Maryland.