Metadata Training in New Mexico: A Case Study

Rebecca Lubas, Ingrid Schneider, Amy S. Jackson

I.   Introduction

Describing cultural heritage materials is a common challenge in the libraries, archives, and museums community. Cultural heritage professionals must be well trained in order to follow best practices for metadata creation. In the current economic climate, having the opportunity to attend training events is challenging for everyone; however, it may be doubly challenging for members of cultural heritage institutions in states like New Mexico. New Mexico is a geographically large state with few large urban centers and a population density of sixteen people per square mile, in which cultural heritage is maintained by many small and geographically remote institutions. In addition to being far from each other, many of these institutions are also far from the closest cities where more training opportunities may exist, such as Denver, Tucson, Phoenix, or Dallas. Nonetheless, the benefits of using common metadata schemes and practices within the cultural heritage community make it imperative to find new solutions for providing training in new standards.

II.   More Technology, More Options, More Need for Training

Improvements in computer and network technology have opened a "wider information landscape" (Riley, Shepherd, 2009). Libraries and other cultural heritage institutions are seeing an increased demand for digital collections. These collections are made up of both digital representations of physical objects and objects that were "born digital." The development of

digital collections is changing the way that cultural heritage professionals look at the process and functionality of description.

The potential of digital collections is only as good as the ability to discover them (Library of Congress Working Group on the Future of Bibliographic Control, 2008). Providing a quality description of digital objects requires different information from that needed to provide a quality description of a physical object. In the mid-1990s, description standards for digital objects began to emerge. These standards were created in a framework based on physical object description and "assumptions about technical architecture, service models and user behaviors [that]…were formed pre-Web" (Dempsey, 2006). Different communities developed distinct standards, a pattern inherited from describing physical objects where different descriptive practices grew around different resource types such as museum items, archival items, and bibliographic items (Library of Congress Working Group on the Future of Bibliographic Control, 2008). In this world of unique physical objects, different descriptive practices were not a problem because information from different institutions types very rarely, if ever, crossed paths. The union catalogs that did exist were generally confined to one type of cultural heritage institution and did not offer interactive, interoperable discoverability.

This tradition has left the cultural heritage community with a number of available standards for digital object description; including Dublin Core (Dublin Core Metadata Initiative, 2008), Encoded Archival Description (EAD) (Library of Congress, 2002), Visual Resources Association Core (VRA Core) (Visual Resources Association, 2007), Categories for the Description of Works of Art (CDWA) (J. Paul Getty Trust and College Art Association, 2006) ,

Cataloging Cultural Objects (CCO) (Baca et al., 2006), and Describing Archives: A Content Standard (DACS) (Society of American Archivists and Hensen, 2004). Additionally, description standards for digital objects are not limited to cultural heritage institutions; the sciences and humanities have also developed specialized standards for digital object description, as have for-profit businesses, such as insurance companies.

The beginning of the last decade saw an increase in search engines' skill "at spidering deep into databases" (Shreeves, Riley, & Milewicz, 2006), and the development of aggregators that brought data from institutions of all disciplines together. When these technologies began interacting with the digital collections of the cultural heritage community, it became apparent that developing standards in a mindset still largely influenced by the world of physical collections had led to digital collections that were highly "siloed" into distinct, non-interacting segments. As the networked world became increasingly flat, the "siloed" nature of the digital collections made them increasingly unwieldy and inconvenient to users. Nonetheless, the spiders and aggregators found ways in, and cultural heritage institutions began "seeing their digital content and metadata showing up everywhere" (Shreeves, Riley, & Milewicz, 2006).

Innovators within the cultural heritage community responded by embracing the new possibilities and finding ways to actively open access to their metadata. In the process, they found that translating information between the disparate description standards was a lengthy and labor-intensive process, and was often not enough to create a quality aggregation of information (Riley, Shephard, 2009). The aggregations that were created were often difficult, if not impossible, to understand outside of the context of the local variations of the description

practices under which the metadata had been created. These difficulties have brought about a movement towards shareable metadata, a core set of elements and best practices that could work in conjunction with the disparate standards to provide enough basic information about a resource to let metadata harvesters successfully interpret the information and users determine their interest in learning more about it. Using an agreed upon "lowest common denominator" metadata scheme, such as Dublin Core, was only the first step in the process. Metadata also needed to be created with shareability and interoperability in mind (Shreeves, Riley, Hagedorn, 2007).

Of course, the widespread, consistent provision of shareable metadata requires a widespread, consistent training of cultural heritage professionals at all levels. This is not new for the community; cultural heritage institutions have always had in common the need for extensive training in order to follow their respective best practices for object description. Traditionally, though, this training has been insular, confined within the respective branches of cultural heritage institutions; and has been a slow, methodical process.

The cornerstone of traditional library metadata training rests with the long-running standard, *Anglo-American Cataloging Rules* (AACR2) (American Library Association, 1998), expressed in MAchine Readable Cataloging (MARC) language. Often, while MLIS librarians may have taken a class in Cataloging and Classification, the real training took place on the job. Cataloging training in libraries usually took the form of a multi-year process that gradually moved the new cataloger to progressively more complicated cataloging, under the tutelage of an experienced cataloger. Sometimes catalogers would be sent to off-site workshops, but these would usually be considered supplements to a more apprentice/journeyman/master type training program. This

method of training has persisted and survived changes in cataloging, such as the move to using a bibliographic utility such as OCLC, the move from card catalogs to electronic catalogs, and the proliferation of digital formats.

Cataloging training in archives and museums has followed a similar model of tutelage. However, archives and museums often have smaller cataloging operations and/or more frequently use institution-specific classification schemes, leaving these institutions with relatively few cataloging experts, both within and outside of the institution, available to train and assist new catalogers. In states like New Mexico, where the nearest expert may be hours away, travel time may also act as barrier to access to expertise.

In today's cultural heritage institutions, there are a number of reasons why this stately approach to description training is no longer practical and sustainable. Ever-decreasing budgets have meant that cultural heritage professionals have to do less with more. The rapid proliferation of resources and vastly changed user expectations resulting from technological improvements have added to what cultural heritage professionals must achieve within in their limited resources, and changed the speed at which they are expected to make resources available. Training programs such as "Metadata for you and me" have shown that new approaches to metadata training can be successful, and that there is a need for these types of opportunities within the community (Riley and Shreeves, 2008).

III. New Mexico Metadata Day

Cultural heritage institutions in New Mexico focus primarily on Hispanic, Native American, and Latin American cultures, as well as Southwestern United States history. The state is anchored by

two major research institutions (University of New Mexico and New Mexico State University), including their libraries and museums, and most other cultural heritage institutions are small, independent organizations, primarily museums. The local museums have small numbers of staff, who have primarily been trained in the museum tradition.

In 2009, the University Libraries at the University of New Mexico organized a project celebrating the centennial of New Mexico Statehood, and invited many of the cultural heritage institutions in New Mexico to participate by providing historic images with descriptive metadata. Participants in the project include the Albuquerque Museum, the New Mexico History Museum, the Indian Pueblo Cultural Center, the Museum of International Folk Art, the National Hispanic Cultural Center, New Mexico State University Library, and others. The platform chosen to host these images was ContentDM, and the project stakeholders agreed on the *Dublin Core Metadata Best Practices* created by the CDP Metadata Working Group. Due to the common subject matter and shared platform, metadata creators trained in different backgrounds needed common training and understanding of metadata fields. The project stakeholders organized training for members of the participating institutions, but the situation made it clear that a community of metadata creators at New Mexico cultural heritage institutions would be advantageous.

With these challenges and needs in mind, two librarians at the University of New Mexico University Libraries and one librarian at the New Mexico State University Library created a collaborative approach to provide metadata training to cultural heritage professionals in the state of New Mexico. The goal was to provide a low-cost metadata training day with local expertise, and to lay the groundwork for the creation of a community of metadata creators in the New

Mexico libraries, archives, and museums (LAM) community. In order to help alleviate the travel burdens on participants, sessions were held both at the University of New Mexico in Albuquerque and New Mexico State University in Las Cruces, located in the central and southern regions of the state, respectively. The sessions were widely advertised, and participation in the training sessions was free and open to anyone. The sessions varied in size, which created very different experiences for both the trainers and participants. A survey was offered to participants to help the presenters design future sessions.

The sessions were advertised on LAM electronic mailing lists, in a regional library newsletter, on Facebook, and via word of mouth. Attendees included professional librarians, library staff, library school students, museum curators, archivists, public librarians, private librarians, and cultural center staff. Attendance for the Albuquerque training session was high, with nearly 50 participants from Albuquerque, Santa Fe, and other near-by areas. The attendance for the Las Cruces session was much lower with only four attendees, but the session drew participants from West Texas and created new relationships in a region has not seen frequent training or collaboration opportunities.

The schedules for the Albuquerque and Las Cruces sessions were slightly different, but both centered around three one-hour lectures covering different aspects of metadata. The first lecture was an overall introduction to the metadata, which included a definition of what constitutes "metadata", examples of metadata from everyday life, and a brief overview of several common metadata schemes. The second lecture presented the elements of qualified Dublin Core as defined by "Using Dublin Core" (Hillman, 2005), and "DCMI Metadata Terms" (DCMI Usage

Board, 2010). The third lecture covered the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), including a basic explanation of how metadata harvesting works and examples of metadata records demonstrating the importance of consistent, shareable metadata.

To help reduce costs for participants, the sessions were scheduled for mid-day in the hopes that participants would be able to drive to and from the sessions that day without having to spend the night locally. Each session began at 10:00 a.m. Because the University of New Mexico is a larger institution with established digital collections and metadata creators, the trainers were able to put together a panel of other metadata creators to discuss the practical application and use of metadata. Also, because many of the participants were from cultural heritage institutions and library divisions working with finding aids and EAD, the Albuquerque session offered an opportunity for an informal discussion on the use of EAD. The session lasted until 5:00 p.m., with a break for lunch and an afternoon refreshment break.

While the Las Cruces session included the same lectures, there were no additional experienced metadata creators for a panel session, and there was no interest in an informal EAD discussion. The last event of the Las Cruces session was the refreshment and socializing break. The event was scheduled to end at 3:30 p.m.

Neither the trainers nor the Las Cruces participants found the small group size and lack of a formal panel discussion to be a disadvantage. The trainers found that the difference in group size changed the dynamics of the session, and actually felt that the Las Cruces session was the more successful of the two. The larger session, in Albuquerque, had a very formal feel and participants

seemed hesitant to ask questions. Also, the sheer number of people taking part made it difficult to sustain one discussion during the break, and many participants broke into smaller groups for networking. The Las Cruces session, on the other hand, was very informal and both the trainers and participants felt more comfortable pursuing further discussion of the topics. There was a great deal of conversation between the trainers and participants both during the lectures and the break, and many of the participants stayed past the official ending time while specific projects and institutional circumstances were discussed.

Several weeks after the sessions, participants were invited via email to complete a survey about their experience at New Mexico Metadata Day. The survey was a mixture of open and closed questions administered through the University of New Mexico's SurveyMonkey account. Participants were informed that their participation in the survey was voluntary, and that the results would be used primarily for the evaluation and improvement of training methods, but that there was a possibility that a summary of the results would be used in publication for the purpose of sharing the experience with the larger cultural heritage community. There were thirteen respondents to the survey invitation, a 26% participation rate.

The survey consisted of eleven questions, including questions on why the participants wanted to attend the sessions, whether employers supported the participant's desire to attend, what the participants' expectations were prior to the sessions, what they had gained by attending the session, the participants' prior knowledge of and experience with non-MARC metadata, what the participants felt were the most and least helpful elements of the sessions, and suggestions for future training sessions.

The majority of respondents believed that the knowledge they gained at the session would be useful either in their present or future job duties, and the majority of employers shared their enthusiasm for attending the session. Many of the respondents learned about metadata in school and through other workshops, and some had experience creating non-MARC metadata. All attended the sessions with the hopes of gaining additional knowledge and understanding of metadata for reasons directly related to professional duties, and the majority of respondents expressed a feeling that they knew more about metadata in cultural heritage institutions after the session than they had prior to the session.

Questions related specifically to the lecture content showed that while the respondents found the lectures informative and helpful, there were strong feelings that a more interactive session using examples from participants' collections would be more useful. Many survey respondents indicated that having the chance to network with other professionals was either something they hoped to gain from the session or was one of the aspects they found most helpful when looking back on the sessions. Some responses implied that the participant was unsure of the level of experience required prior to the session, indicating that the invitations for future sessions should better clarify the material to be covered and expected prior knowledge.

While thirteen responses to a feedback survey is a very small sample to base hard conclusions on, the responses were valuable to trainers in terms of improving their lectures. After the initial two sessions of New Mexico Metadata Day, the trainers had the opportunity to give the workshop two more times to two very different audiences.   Each instance informed the next.

- Why were you interested in attending Metadata Day?
- Do you agree with the following statement: The knowledge you gained at Metadata Day will be useful in your present or future job duties.
- Did your employer share your interest in attending Metadata Day?
- What was your experience with metadata prior to the training session?
- Have you previously created metadata in a schema other than MARC/AACR2? If so, which schema have you used?
- What did you hope to gain from attending Metadata Day?
- What did you find to be the most interesting or helpful part of the workshop?
- What did you find to be the least interesting or helpful part of the workshop?
- How do you feel the workshop can be improved?
- Do you feel that you know more about metadata in cultural heritage institutions now?
- Do you have any suggestions for future metadata training workshops?

Figure 1. New Mexico Metadata Day Feedback Survey Questions

IV.   Improvements and Lessons Learned

The initial offering of the workshop in and of itself was a success. Many of the attendees from the smaller institutions expressed appreciation for an opportunity to learn anything about metadata. Having multiple presenters allowed for different views and opinions to be shared. During the lectures a presenter often involved the other presenters in audience questions for additional perspectives.

As with any training opportunity, there was definitely room for improvement and refinement. A large portion of the day was devoted to lectures and straightforward presentations. Based on the survey feedback, presenters felt that increasing the amount of interactivity would help keep the audience engaged in the materials. The participants seemed more engaged after asking them to

evaluate metadata records, and more examples were added in subsequent iterations of the workshop. Additionally, audience members from outside of the library discipline commented that there was too much library-specific jargon, and presenters worked to eliminate this jargon in future presentations.

The shareable metadata and OAI-PMH lecture was intended to demonstrate to participants how metadata was used outside of the local context, and how metadata is viewed by service providers. The contextualization of OAI-PMH and tips for creating shareable metadata seemed to make sense to users, but discussion of the hands-on harvesting of metadata seemed to overwhelm members of the audience that did not have much experience with technology. The amount of technical detail needed to be tailored to the audience.

In the spring of 2011, the presenters had the opportunity to present the session at the state library association conference as a half-day preconference workshop. The conference draws on a variety of librarians from public, school, and academic institutions. The presenters used feedback and their own observations from the original sessions to refine the materials for this session.

The introductory lecture cut some of the examples of metadata languages other than Dublin Core to avoid overwhelming the audience with superficial information. The Dublin Core lecture focused on the basic DC elements, rather than attempting to cover qualified DC in detail. The slides were altered to present basic information, rather than exhaustive information, to encourage active listening. More digitized photos were included to show real-life examples. The technical aspects of the OAI-PMH lecture were downplayed, and the hands-on portion of harvesting from

a data provider was eliminated entirely. New examples of metadata records with questionable uses of metadata fields were included, and the participants were asked to evaluate the records based on shareable characteristics. Participants commented that they enjoyed the interactive portion of the session, and the presenter felt that due to less emphasis on technical details she was able to keep the audience more engaged. No questions arose about jargon words or terms.

Following the state library association presentation, the presenters were invited to present to personnel from the University of New Mexico's information technology (IT) department. The IT department had been hearing the term "metadata" mentioned often in the past several years in context of web authoring, and they wanted to learn more about it. They understood that the presenters would be speaking from the library community's point of view and would provide less technical detail and more of an overview of metadata and issues around metadata. The presenters took into account the different background this audience had than the audience they normally presented to, and they tried to minimize library jargon. The introductory lecture especially took the opportunity to explain library context, such as the development of the MARC tagging language and early development of integrated library systems that had long-ranging impact on metadata practice. The presenter giving the OAI-PMH overview was able to fully explain all XML examples and spend more time on the protocol itself than in previous presentations. The IT department said that they enjoyed the "geeky" part of the OAI presentation, and were surprised overall at the lack of defined content standards for the use of Dublin Core. At the end of the session, the presenters and the IT personnel acknowledged a common ground in the desire for consistent data and clearly defined standards.

V. Conclusions

According to a study of adult learning perceptions, "adults expect learner-centered settings where they can set their own goals and organize their own learning around their present life needs" (Donaldson, Flannery, and Ross-Gordon, 1993, p. 148). The varying needs of librarians, cultural heritage institution professionals, and other information professionals that the presenters encountered in their various iterations of the Metadata Day workshop upheld this conclusion.

The presenters recommend keeping the following points in mind when designing or adapting training in metadata practices:

- Understand the tradition in which your audience members are trained
- Understand the audience's familiarity with vocabulary terms
- Create interactive examples to help keep the audience focused and engaged
- Customize technical details in the lectures to fit the audience.

Creating a community with a sense of shared mission does not happen in one training session. Providing the session multiple times helps identify members of the community interested in metadata, and has provided more opportunities to present to new audiences. More introductory presentations will help grow the base of metadata creators, but more in-depth workshops will help bring these audiences together again. The presenters hope that the New Mexico Metadata Day sessions will help create a shared, unified vision of descriptive metadata in cultural heritage institutions in the state of New Mexico.

References

American Library Association. 1998. *Ango-American cataloguing rules*. Chicago, Ill: American Library Association.

Baca, Murtha, Patricia Harpring, Elisa Lanzi, Linda McRae, and Ann Baird Whiteside. 2006. *Cataloging cultural objects: a guide to describing cultural works and their images*. Chicago: American Library Association.

CDP Metadata Working Group 2006. *Dublin Core Metadata Best Practices, Version 2.1.1*. Available at http://www.mndigital.org/digitizing/standards/metadata.pdf (retrieved June 7, 2011).

DCMI Usage Board. 2010. "DCMI Metadata Terms". Available at http://dublincore.org/documents/dcmi-terms/ (retrieved June 21, 2011).

Dempsey, Lorcan. 2006. *The (digital) library environment ten years after*. Bath, U.K.: UKOLN.

Donaldson, Joe F., Flannery, and Ross-Gordon. 1993. "A Triangulated Study Comparing Adult College Students' Perceptions of Effective Teaching with Those of Traditional Students". *Continuing Higher Education Review*. 57 (3): 147-65.

Dublin Core Metadata Initiative. 2008. *Dublin Core Metadata Element Set, Version 1.1*. Available at http://dublincore.org/documents/dces (retrieved June 7, 2011).

Hillman, Diane. 2005. "Using Dublin Core". Available at http://dublincore.org/documents/usageguide/ (retrieved June 21, 2011).

J. Paul Getty Trust and College Art Association. 2006. *Categories for the Description of Works of Art*. Available at http://www.getty.edu/research/conducting_research/standards/cdwa/index.html (retrieved June 7, 2011).

Library of Congress. 2002. *EAD: Encoded Archival Description: Version 2002 Official Site*. Available at http://www.loc.gov/ead (retrieved June 7, 2011).

Library of Congress. 2008. *On the record report of the Library of Congress Working Group on the Future of Bibliographic Control*. Washington, D.C.: Library of Congress. http://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf (retrieved June 21, 2011).

Riley, Jenn, and Shreeves, Sarah L. 2008. *Metadata for You and Me: Current and Emerging Trends in Metadata and Content Sharing. Presentation for Digital Library Federation Fall Forum 2008*. Available at http://hdl.handle.net/2142/9158 (retrieved June 21, 2011).

Riley, Jenn, and Kelcy Shepherd. 2009. "A Brave New World: Archivists and Shareable Descriptive Metadata". *The American Archivist.* 72 (1): 91.

Shreeves, Sarah L., Jenn Riley, and Kat Hagedorn. 2007. *Best practices for OAI PMH data provider implementations and shareable metadata: DLF/NSDL working group on OAI PMH best practices*. Washington, D.C.: Digital Library Federation.

Shreeves, Sarah L., Riley, Jenn, and Milewicz, Liz. 2006. *Moving towards shareable metadata*. University of Illinois at Chicago. Available at http://firstmonday.org/issues/issue11_8/shreeves/index.html (retrieved June 21, 2011).

Society of American Archivists, and Steven L. Hensen. 2004. *Describing archives: a content standard*. Chicago: Society of American Archivists.

Visual Resources Association. 2007. *VRA Core 4.0*. Available at http://www.vraweb.org/projects/vracore4 (retrieved June 7, 2011).