

Running Head: INDEPENDENT STUDY

Independent Study:

Metadata, Art Images and Digital Asset Management at Acme College

April Younglove

Emporia State University

Abstract

This is a summary and analysis of a library science student's experience using ContentDM and Excel to assign metadata to digitized art slides. The Independent Study project described was preformed for Acme College from January to April of 2008. This essay will include a description of the project and it will attempt to place the project into the context of the larger academic and library science community. It will also establish links between the metadata work done for Acme College's Library and the student's previously completed library science coursework.

Background

For my Independent Study, I began volunteering as a Metadata Specialist for Acme's College's Digital Assets Management (DAM) Team in January of 2008. I have committed to working with Acme's Library in Washington to continue to digitize their art slides until August of 2008, when the project is scheduled for review and possible completion. P, Technical Services Librarian for Acme Library is serving as the project supervisor, and R, a library and archives paraprofessional from X Library, is contracting with Acme as the DAM Team's primary consultant and project engineer. Additional members of the DAM Team include L. and S. of Acme College's art department, and P., the library's IT Specialist.

The goals of the DAM Team are to:

- Develop a centrally-managed, flexible, and sustainable collection of digital images and related materials;
- Integrate digital materials into the curriculum;
- Provide good in-class presentation tools and easy student and faculty access outside of class;
- Develop methods for integrating with other content resources, like ARTstor;
- Focus on art but develop systems and processes that can be used across the curriculum. (Smith, 2008, n.p.)

S. originally recommend that Acme pursue digitizing its art slides, in part because the database ArtStor does not adequately meet her needs, and in part because she wishes to have the added flexibility of using the images in her PowerPoints. It is very easy to create a PowerPoint slide that shows two images side-by-side, but very difficult to do this with traditional slide projectors. Another reason for Acme College to invest in slide digitization at this time is that traditional slide projectors are becoming more difficult to obtain and find parts for. In 2004, Kodak, the largest supplier of slide projectors, made the decision to discontinue its projector manufacturing (Kodak, 2004, n.p.). Although art history professors, who have traditionally used art slides to present their lectures, were already beginning to experiment with using formats such as PowerPoint to deliver content, Kodak's announcement came as a clarion call for all colleges to "go digital."

Stakeholders

The purpose of the DAM Team is to serve the interests of the Acme College community. Acme College is a public college with two campuses and a student body of around 13,000. Originally a two-year junior college only, Acme now offers bachelor's degrees in Technology, Dental Hygiene and Social Work and a master's degree in Social Work from Y University at Acme College (Acme College, n.d.). The DAM Team's goals place its digitization project into the larger framework of the college's master plan, especially the Master plan's objectives to include better integration of technology into Acme's classrooms, and provide

better support for e-learning in the course content management system, Blackboard (Acme College, 2007).

Stakeholders in the project include Acme's present and future students of art history, its art history instructors, and Acme library staff. In the future, after the digitization of art slides is complete, it is hoped that the ContentDM platform can be expanded to benefit students and teachers in other disciplines outside of art history. Funding for the current project has been provided through a grant secured by the college's art faculty from Acme College Foundation, a private charity created by Acme College alumnus. A monetary gift was also given from the Foundation to Acme Library for art slide digitization. The success of the project will improve the classroom experience for students and staff at Acme College, which will benefit both the college and the Foundation.

DAM Communities of Practice

Acme College's slide digitization project is not pioneering in the field of Digital Assets Management (DAM), but follows in the footsteps of other forward thinking libraries and archives across the nation. For the past several years, similar efforts to digitize slide collections and make electronic content more available in the classroom have been and are underway in hundreds of American colleges and universities. In the book *Digital Libraries: Integrating Content and Systems* by Dahl, Banerjee and Spalti (2006), Digital Assets Management (DAM) is defined as an institution's commitment to providing access to and preservation for its digital materials. Some examples of these types of materials include:

“faculty research publications; digitized images from historical photograph collections; university press or other institutional publications; rare book and manuscripts; . . . [and] art slide collections” (p.114). In 2004, over a dozen case studies of emerging DAM work were published in the book *Metadata in Practice* by Diane Hillmann and Elaine Westbrooks.

As a smaller college, with a high concentration of first and second year undergraduate students seeking associate’s degrees or preparing for transfer, Acme looked to established higher learning institutions in the area that have completed similar slide digitization work. Acme College chose to use CONTENTdm, created by the nearby University of Washington, for its management software (Dahl, Banerjee and Spalti, 2006, p.124). It also made a conscious attempt to mirror cataloging decisions made by Reed College’s slide digitization team. Acme is employing the same 20 VRA Core categories and Getty Vocabulary that Reed College is using. This was a way for Acme to jumpstart their project by building on Reed’s experience and authority.

Unfortunately, even though many academic institutions are doing slide-scanning projects, they are not able to reduce costs by sharing their finished results with one another. Libraries have a strong tradition of sharing finished catalog records with one another to reduce labor duplication. However, in this case, although the slides as physical objects are owned the institutions, P. explained to me that many of their digital images may not be sharable due to copyright concerns. Most image collections made from slides, including the

compilation I am currently helping to create for Acme, are not publicly available online and are accessible only to those who are given permission through the institution.

When comparing Acme's DAM efforts with other nearby colleges, it is important to keep in mind that Acme has a nontraditional student body and fewer monetary resources than the average college. The majority of Acme's slide digitization grant money will go to equipment costs and to CONTENTdm for storage space and management fees ([*2008/09-2009/10 Budget for DAM Digitization Project*], n.p. 2008). While most digitization projects can afford to at least two FTE staff as well as a few work study students for daily project tasks, Acme is relying on one part-time paraprofessional consultant (R.) and one less than part-time volunteer (myself) for all of its slide scanning and metadata entry, which severely limits the amount of time it can devote to project development and research.

Project Details

Once a week, for 5-8 hours at a time, I have been traveling to Acme College's Acme library to examine their slide images and enter metadata for those images into an Excel spreadsheet. My task is to help the DAM Team assign data to the 10,000 art slides that are already owned by the institution for later uploading into the management software, ContentDM. There are 20 possible fields to enter information into. Three of the 20 fields (Type, Format and Rights) are static and do not require individualized data. Of the remaining 17, the

most important fields to enter data into are title, creator, and date. However, only the title field is required and a record may be considered complete even without a creator or date if none can be located.

Before I assign metadata to the images, R. accesses the slides that the art faculty deliver to us. Recently, he has also scanned in images from the textbook *Gardner's Art Through the Ages* (Kleiner, F. S. & Mamiya, C. J., 2004) at S.'s request. R. assigns each image an accession number and scans it into a JPG2 format for sharing through CONTENTdm and a lossless TIFF format for archiving in an on-site hard drive. R. then either enters metadata for the slides into Excel himself or sets the slides aside for me to add the metadata later.

Because the slides are organized in accession order, slides with correct labels are mixed with slides that are poorly labeled and some slides are duplicates of one another. The college has amassed its slide collection over a long period and instructors through time have chosen to write or print labels for the slides with differing levels of accuracy and completeness. Because many slides are unlabeled or were labeled incorrectly, I cross-reference what is written on each slide with what I know personally from my courses as an art history minor, with the *Gardner's Art Through the Ages* textbook (where many of the images came from originally), and occasionally with internet resources such as the art object's owner's website (typically a large museum).

There are a number of issues that are unique to cataloging images. When assigning metadata to more traditional media (eg. books, videos, periodicals),

authoritative information about the object can be gleaned either from text printed on the object or by examining and measuring the object itself (counting its pages, for instance). However, with artwork, and especially with pictures of artworks that are once removed from the actual object, the task becomes trickier. If one has a scanned image of a slide that was taken from a textbook illustration of a model of a building in Greece, what is the artwork and who created it? Is it the textbook photograph and photographer? The model and its builder? The building upon which the model was based and its architect? If one is faced with an unlabeled image of an artwork that one knows nothing about, how does one discover what its title is? For a good discussion of these and other image cataloging dilemmas, see "Getting the Picture: Observations from the Library of Congress on Providing Online Access to Pictorial Images," by Caroline R. Mann (1999). Although not all of the questions surrounding the cataloging of images have been resolved, and new ones are continuing to emerge as digitization continues, there are some standards that help guide professionals in making these types of decisions.

At the outset of our art slide scanning task, P. provided me with a list of three resources to be used as standards guides for our project: a link to the Getty Thesauri, a link to the Visual Resource Association's (VRA) Core standard and a link to VRA's Cataloging Cultural Objects (CCO) guidelines. The Getty Thesauri are controlled vocabularies for art and architectural keywords, artist names and geographic names to be entered as specific data values. For our project, we use eventually began using both Getty's The Art & Architecture Thesaurus (AAT) and its The Union List of Artist Names (ULAN). The VRA Core, meanwhile, is a data

element standard guide (it determines the above listed 20 categories). VRA Core was created by the Visual Resources Association (VRA), a professional organization devoted to the science of image management and preservation. CCO's guidelines, also created by VRA, are currently the chief authority regarding metadata content standards for cataloging physical and digital images. Although many of the "rules" of CCO are framed as "recommendations," the guide comes from a decidedly archival tradition and emphasizes hierarchical relationships, establishing relationships between records, the ability to create collection and series level records, the need for authority headings, and valuing consistency over specificity.

Our use of the above standards has varied over the life of the project. Because there were many uncertainties about how closely I was supposed to follow the standards and when it would be appropriate to do so when we first began, I started a shared log of our decisions, both to communicate my practices with R. (with whom I work most closely on this project) and to read about any precedents that R. would decide to set while entering metadata.

Initially, I attempted to use the AAT Getty Thesaurus for as many fields as possible, but it was later decided that in the interests of time, and because it would not be appropriate for our audience, that we would not use the AAT. At that time, we experimented with using the simpler controlled vocabulary that comes with the CONTENTdm package instead. However, because this vocabulary is only available within the subject field as a drop down option in the

CONTENTdm program and because data is first entered in Excel, it was decided that all subject fields will be left blank until a later date. R. proposed that in the future, when the project is nearing completion, we might be able to go back and fill in some of the subject fields at our leisure. However, with multiple people entering data it became clear that consistency was becoming a problem. Finally, R., P., and I agreed that it would be better to use some controlled vocabulary. We decided to use AAT and ULAN for three select fields: creator, period.movement and period.style. We also decided to standardize the measurement field so that artwork dimensions are listed only in centimeters. These decisions remain the current project practices regarding controlled vocabulary.

The 20 VRA Core categories that we have decided upon have not changed over the course of the project. All 20 fields are able to crosswalk into the more transferable Document Type Definition, Dublin Core, in case of future export. CONTENTdm allows users to enter fields into a form view and does not require users to generate the tags that appear in a RAW mark-up view. Therefore, it was not necessary for me to learn how to format the VRA tags. Instead, I used an Excel template with pre-labeled Core categories and entered metadata into the appropriate fields in plain text.

CCO includes standards for database construction as well as for how to format data content. However, only the former was relevant to this project as we used CONTENTdm and were not customizing the content management system to meet CCO standards or building a database from scratch. Unfortunately this

means that I was not able to create any authority records (to standardize museum location names, for instance), or to create relational records, or links between records in the way that CCO suggests. CCO Guidelines were helpful though in answering some of the above trickier cataloging problems and allowing us to make choices about how to enter data. For instance, in the above mentioned example about the slide that could be either of a model or of a building, CCO privileges the original work over the image of the work. Therefore, one catalogs the data about the original object's format: architecture. The only drawback to using CCO for this purpose is that it was originally designed for an environment where the real objects are present or accessible to the catalogers and not with the idea that it would be used for an image only database. Its broad guidelines of determining intrinsic vs. extrinsic qualities, valuing consistency over specificity, and remaining object-centered, can help the cataloger make thoughtful and reasonable choices though.

Currently, as of late April 2008, I am at the mid-point of the slide-scanning project's original timeline. We have completed the scanning and metadata for approximately 1,500 of a total 10,000 slides. At present, I am editing the information for the first 1,500 images in CONTENTdm so that all records are consistent and reflect our current data standards. This is a slow process as CONTENTdm retrieves and displays search requests for multiple files in alphabetical order by title, making it impossible to batch edit by accession number. Therefore, in order to edit image entries, each record must be searched for and edited individually. In addition to editing past slides to reflect current

standards, another way for us to further standardize metadata at this point, R. has recently suggested, is for me to do all of the future metadata entries instead of sharing the task between the two of us. However, this idea would further delay our rate of progress.

As we are cataloging, art instructor bring in new materials for us to add to those we already have queued for scanning, typically because they need to use those materials in the near future and would otherwise scan those images at home, which is a redundancy of labor. P.'s impending retirement from her position and the recent hiring of a new director create an environment of transition for our project and may also impact our timeline.

Connection to coursework

I am currently two terms away from graduating from Emporia University with a Master's Degree in Library Science. Because I have completed such a large percentage of my coursework at this point, I have been able to find many ways to apply the theories, methods, and knowledge that I have gained from my classroom learning to my independent study experience. There are articles and textbooks from many of my different classes that come to mind as I work at Acme. Some courses in their entirety have prepared me for my Independent Study project. I can also recall specific projects and group work experiences that have helped expose me to working as a part of a team on a long-term project. Although I would like write a detailed explanation of how each course that I have taken with Emporia has aided me on my journey thus far, unfortunately I do not

have space here to do the topic justice. I have chosen, however, to highlight the connections between eight different courses and my independent project work.

In my first term at Emporia, Summer 2006, I took *Foundations of Information Management*, and *Information Transfer in the Disciplines*. In *Foundations of Information Management* with Dr. Linda Lillard, Lillard introduced us to librarianship as a community of practice with professional ethics. She strongly encouraged us to join professional organizations and become involved in the library world through volunteer work. This recommendation is what inspired me to become involved with library organizations and led me to apply for different volunteer jobs like the one I am presently doing for Acme College. *Information Transfer in the Disciplines* introduced me to the concepts of paradigm shifts and the information transfer cycle. It is clear to me, especially with the research that I have done for this independent study, that libraries are undergoing a paradigm shift from knowing how to manage physical collections to learning how to manage digital collections of items born digitally or transferred into a digital medium.

My second term included the classes, *Theoretical Foundations of Service: Diagnosis and Customization*, and *Organization Theories for Administering Information Agencies*. In *Diagnosis and Customization*, I learned how to do an information needs assessment and had my first experience as an adult working on a project with a group. I recalled my information needs assessment work from this class when meeting with Acme's art faculty members. When those of us with library science backgrounds did a second interview with the art instructors to

learn about what they desired in the finished slide database, it became clear that we made too many assumptions after their first interview. For instance, early on in the project, we discussed points of access with the instructors. They felt that they needed very few, just title and date. However, in a later conversation, one of the art teachers made reference to how she would be able to look up the term “women artists” in order to find images by or about women in our database. At this point, it became clear that there was an information gap. It would not be possible for them to access images by women artists unless we entered that information into one of the record fields, and neither the record field title nor date specify whether the artist is a man or a woman. Working on a complex group project with five other people, performing joint research and creating multiple coordinated presentations in *Theoretical Foundations of Service* also prepared me for a real team environment where different group members have different interpretations of what the project is and how it should be completed. In my current project, I use similar listening and negotiation skills.

One of the instructional objectives for *Organization Theories for Administering Information Agencies*, was “Knowledge of the different factors that shape an organization, including its present state, its planning process, and its strategies and goals.” In the book *Organization Theory* by Mary Jo Hatch (1997), assigned in that class, Hatch argues that, “you will want to know the theories that others are using and how to create your own so that you can more consciously (and conscientiously) participate in these processes” (p.6). For instance, it is good to be aware that a company that perceives itself as mechanistic, that is, as

stable with specialized parts, will be averse to change while an organization that is structured organically will likely be decentralized and require employees to have great flexibility (Hatch, 1997, p.76). After thinking about what kind of organization Acme College represented, I determined that it is bureaucratic, which is a traditional college organizational structure. Realizing this allowed me to better understand and accept the slow and deliberate pace of our project to date.

Like *Organization Theory*, the texts, *Digital Libraries: Integrating Content and Systems* by Dahl, Banerjee and Spalti (2006) and *Metadata in Practice* by Diane Hillmann and Elaine Westbrooks (2004), for the course *Metadata Frameworks* from Spring of 2007 are particularly relevant to my Independent Project. I have consulted both works for examples and explanations of similar projects while working with Acme. In my *Metadata* class, the instructor, Terry Reese, introduced me to a set of terms and concepts that I use in my workplace and at my volunteer job almost constantly. Reese is an expert in metadata crosswalking, especially for images (because of his involvement with the NorthWest Digital Archives). Even though I did not think at the time that lessons on image management were relevant to me, his class put me in an excellent position for understanding this current project. *Metadata Frameworks* familiarized me with creating a Dublin Core record from scratch, how to crosswalk data from one metadata standard to another, and how DAM systems differ from one another. My main regret is that I currently am not able to be involved with more technical aspects of Acme's digitization project (for instance, creating the

crosswalks from VRA Core to Dublin Core, or working more directly with CONTENTdm).

This past term, I have been taking *Introduction to Archives* and *Theory of the Organization of Information* while doing my independent study work. I have used both of these classes to help inform my independent study, taking advantage of projects and presentations to explore the issues that surround image archiving. In *Archives*, instructor Mary Jo Pugh taught me that although the disciplines of library science and archiving overlap, they draw on very different organizational premises. In Dupont's (2007) article "Libraries, Archives, and Museums in the Twenty-First Century: Intersecting Missions, Converging Futures?" the author concludes after an interdisciplinary conference on the title's topic, that:

[T]he relationships among libraries, archives, and museums as institutions depend upon the ways they relate to the collections or content they manage as well as upon the ways they relate to their audiences. These relationships imply certain values, and values, in turn, shape practices. Awareness of differences in values and practices can help library, archives, and museum professionals understand one another better and collaborate more effectively over issues and initiatives of mutual concern (p.14).

In working with R., it was helpful for me to understand that our perspectives would be different since his background and expertise come not from a Library and Information Science education, but rather from an advanced degree in

history and experience with archives. As a cataloger and library science student, I often tend to emphasize the importance of access to information on an item level (which libraries usually provide) rather than at a collection level (which archives usually provide).

For *Theory of the Organization of Information*, I attempted to tie my coursework to my independent study by focusing my group work and my bibliography assignment on related topics. Because I myself was just learning about VRA Core at the outset of this class, I suggested that my group do our presentation about VRA Core. Thus, I had a chance to thoroughly research VRA and to lead a class discussion about VRA Core Categories. I also created an exercise to demonstrate the differences between a controlled vocabulary, a data content standard, and a data element standard for my classmates. For my final bibliography, I chose the topic “Organization of Materials in Museums,” and this gave me a chance to explore image management in a museum context and to compare and contrast those methods with Acme’s methods for image management.

Conclusion

The process of writing this paper has reinforced what I have learned, both over the course of my independent study, and throughout my degree program. It allowed me to think about my work in its larger community context, relating my task to the larger community of financial and academic stakeholders. I have also realized that I am constantly incorporating my in-classroom learning into my current work, whether I am conscious of that fact as I am doing so or not.

Looking back on the work that I have done so far, there have been both pleasurable and frustrating aspects to working on Acme College's slide digitization project. I have tremendously enjoyed viewing and working with art images. Entering in metadata for art slides has refreshed my art knowledge and reinforced my desire to work in some way with the subject of art or with images or artworks in the future. The treasure hunt aspect to tracking down and locating the correct information for a particularly obscure or mislabeled slide is exciting and rewarding. It was fascinating to discover, for instance, that the current location for one particular artwork is "reported as stolen by the LA police department."

DAM Team members and Acme staff have been very warm and accepting, often letting me take over technical services' counter space or not minding that I repeatedly borrow Acme's reference copy of *Gardner's* off reserve for a moment. In my paid day job, I work in a small medical-academic library. Being in a larger technical services space for my independent project, among many other technical staff members, can be very collegial and allows for easy mutual collaboration. I have discovered, though, that I prefer the variety of my day job, where I typically combine ILL tasks, cataloging, and light reference in a single shift over the sameness of only cataloging. In the future, I also hope to attain a job that has more of a leadership role, and I look forward to taking the course *Management of Information Agencies*, which I think will help me achieve this goal.

If I had this project to do over again, I would be more assertive in the formative stages of our work. I would ensure that data would not need later re-editing by asking for more initial shared agreements and by developing a very specific written plan to adhere to. In the initial stages of the project, I did not feel that I had the necessary experience or authority to formulate and present such a plan. However, in the future, I will pursue the creation of a project management chart, even if it feels like it might be impolite or tedious to discuss it with my fellow team-mates.

Taking more time to establish a sound philosophy about cataloging, perhaps by using a particular method or author as an authority, would also have helped clarify the situation. For instance, we could have agreed to draw on the work of Donna Slawsky (2007), an experienced public and private consultant who builds metadata models and taxonomies for Digital Asset Management (DAM) systems and web sites. As she writes in her article "Building a keyword library for description of visual assets," it is important for DAM personnel to employ controlled vocabulary for increased accessibility. This is especially true for visual objects that cannot make up for bad indexing with good full text searching.

Currently, I do not predict that we will meet our goal of 10,000 slides scanned and uploaded with metadata by our initially stated date of August 2008. However, at a DAM all-member team meeting on April 21st, with the new director in attendance, the possibility of extending the grant money into future years was discussed. Plans were made to request future grant monies and to continue the

project on even after I am scheduled to depart in August. This encouraged me greatly to know that the work I have done for my independent study will help Acme College's students and instructors in the future.

References

- [2008/09-2009/10 Budget for DAM Digitization Project]. (2008). Unpublished Acme College document.
- Arms, Caroline R. "Getting the Picture: Observations from the Library of Congress on Providing Online Access to Pictorial Images," *Library Trends* 48: 2 (Fall 1999): 379-409. Also available online at <http://memory.loc.gov/ammem/techdocs/libt1999/libt1999.html>
- Acme College. (2007). *Acme College facilities master plan*. Retrieved April 30, 2008, from http://www.Acme.edu/about_Acme/master_plan/documents/f_master_plan.pdf
- Acme College. (n.d.) *About Acme: College History*. Retrieved April 24, 2008 from http://www.Acme.edu/about_Acme/profile/history.php
- Dahl, M., Banerjee, K. & Spalti, M. (2006). *Digital Libraries: Integrating Content and Systems*. Oxford: Chandos.
- Dupont, Christian. (2007). Libraries, Archives, and Museums in the Twenty-First Century: Intersecting Missions, Converging Futures? *RBM*, 8(1):13-19.
- Giesecke, J. & McNeil, B. (2005a). Teamwork and group dynamics. In *Fundamentals of Library Supervision* (pp. 44-58). Chicago: American Library Association.
- Hatch, M. J. (2006). *Organization Theory: Modern, Symbolic, and Postmodern*

- Perspectives* (2nd ed.). USA: Oxford University Press.
- Hillmann, D., & Westbrook, E. (2004). *Metadata in Practice*. Chicago: American Library Association.
- Kleiner, F. S. & Mamiya, C. J. (2004). *Gardner's Art Through the Ages* (12th ed.). CA: Wadsworth.
- Kodak. (November 2004). *KODAK Slide Projectors: Manufacturing discontinued, October 2004*. Retrieved February 2008 from <http://slideprojector.kodak.com>
- Richman, L. (2002). *Project management: Step-by-step*. New York: Amacom
- Schachter, D. (2004). Managing your library's technology projects. *Information Outlook*, 8(12), p. 10-2.
- Slawsky, Donna. (2007). Building a keyword library for description of visual assets: Thesaurus basics. *Journal of Digital Asset Management*, 3(3): 130-138(9).
- P. (2008). *Digital assets project goals: Integrating digital collections into the curriculum at Acme College*. Unpublished Acme College document.
- Visual Resources Association. (2006). *Cataloging cultural objects: A guide to describing cultural works and their images*. Chicago: American Library Association.