PACSCL/CLIR "Hidden Collections" Project

Holly Mengel's remarks at RBMS:

I am going to talk about how this specific project has fit into PACSCL's consortial efforts and the challenges we have faced, the solutions we have developed over the course of the project, and the benefits reaped because we are working together rather than alone. We feel that these benefits extend beyond the project participants to other archivists considering minimal processing projects who will hopefully learn from what we have done (good and bad). More than anything, we hope that scholars and students use these collections to further their own and the public's knowledge of our collective history.

The project is an ambitious one, reflected by the number of times I have died my hair to cover the grays since starting this job last July. It has been worth every gray—these collections are absolutely amazing and truly deserving of being revealed to the researching public.

 [slides 2-4]

[slide 5]. I would like to thank Christine DiBella, John Armstrong and Jenny Barr for their amazing work . There are two main reasons that the Survey Project has been helpful to us in our work: [slide 5]

[Slide 6] Challenges PACSCL (and almost all repositories) face:

[Slide 7] Backlog is the easiest challenge to see--usually there is little we can do to hide it. What I have learned over the last year is that everyone has it and it is indicative NOT of what is left undone, but how much great stuff is in the area and the efforts repositories have made to save this amazing material from trash cans, dumpsters, and in the best case, basements and attics. However, just because it is saved doesn't mean it is accessible. That is where this project comes in--we will be alleviating only the tiniest fraction of the backlog. However, we hope that by sharing what we learn, what works and what doesn't, repositories will be able to adapt, to a certain degree, our workflows and systems in order to continue work on backlogs.

[Slide 8] Lack of Standardization is another challenge that PACSCL and the archival community at large faces.

Technology is almost never standardized … ranging from repositories with many computers to those who need us to bring laptops because there is not a spare workstation. IT staff also ranges from entire departments to non-existent. Some repositories have been using AT, but for the most part, we are installing AT just before the start of processing. A challenge within a challenge is the question of sustainability—who takes care of potential problems with the software when the repository does not have an IT staff? What happens when we add yet another task to an already overburdened librarian or archivist who is the only person in their department?

[Slide 9] Authority control is probably one of the biggest reasons for all my gray hair …
because we are working across 23 repositories and trying to create a central EAD
repository, the authorities in this project are an essential component of each finding aid.
The importance of a really standardized list of authorities became blatantly obvious when
I realized that the amazing site created by Penn uses faceted searching. We are only
using authorized terms (with the exception of people's names), but there is still a lot of
room for problems … do you use Native Americans, Indians of North America or
American Indians—all of which are authorized terms . If some repositories use one and
others use another, researchers will not necessarily find all the collections relating to their
topic—which defeats the purpose of the central EAD repository.

[Slide 10] In the case of this project, the lack of standardization of technology, authority
control and processing practices had to be resolved, to as large a degree as humanly
possible, because of the centralized EAD finding aid site. Not only do the finding aids
have to import into the Archivists' Toolkit, but we would like for them to look good too.
However, this is no easy task … we know that repositories have been processing and
creating finding aids a certain way for a long time … sometimes, in Philadelphia, a
REALLY LONG TIME. Our goal is to create finding aids that fit, as much as possible,
with the repository's existing guides, but also include all of the essential components
required by the project.

[Slide 11] If we wanted this project to work at all, we knew that we would have to be a
little creative.

[Slide 12] So, how do we jump all these hurdles? Well, I was lucky in that before I
started, the PACSCL group did some of the dirty work … they determined that the
project would employ minimal processing, the Archivists' Toolkit, DACS, and obtain
authorized subject headings from the Art and Architecture Thesaurus and the Library of
Congress.

[Slide 13] Available time and money ensured that we would be minimally processing the
collections. In fact, we are *really* minimally processing the collections if you look at the
time factor … two hours per linear foot for pre-20th century, non-institutional records.
Before the project really started, (and I can hardly remember such a time in my life), I
was not really sure what minimal processing meant for these types of records … and I
certainly did not have even a clue what could be accomplished in the time frame allotted.
I also knew that minimal processing is controversial and could potentially be a problem
within the PACSCL community. (slide)

[Slide 14] So, I made manuals … both a minimal processing manual which addresses:
(slide)

[Slide 15] and an Archivists' Toolkit manual which focuses primarily on what the
students need for this project. Their manual includes (1st part of slide), but does not
include adding authorities or other administrative information which is completed by the

project archivist or myself prior to the students arriving to process. It also includes instructions for proofing and editing work, since our goal is good quality finding aids, and our students are working at an extremely rapid rate. As is so often the case with consortial work, I borrowed from a lot of experts to create the manuals… those in and outside of PACSCL. The Standards Committee, a small group of PACSCL folk focused primarily on this project, met and discussed the big and the small questions that needed to be addressed and proofed my work.

[Slide 16] It seemed like we had a structure in place and then, I decided I need to try it out.
The Drexel University College of Medicine folk offered me a collection and, following the manual, I physically processed it and put the container list in AT in less than 2 hours per linear foot. I was feeling very pleased … until I started writing the biographical note and the scope and content note at which point I discovered

[Slide 17] how fast 2 hours per linear foot really is.

[Slide 18] At the end of that day, I was faced with a frightening discovery: it is impossible to achieve the quality of work we want within the time frame allotted. I felt strongly that brand new student processors could not be asked to do what I could not do after a fair number of years of experience.

[Slide 19] (slide text) So, it seemed necessary to develop a structure of providing information so that the 2 hours per linear foot allotted to the students could be used for processing, not discovering what a collection contains or is about.

I determined that the project would have to create processing plans for each collection that our student processors process. These processing plans include:

> Proposed plan for processing the collection with information from the survey and from the person creating the processing plan;
> Proposed list of series;
> Preliminary biographical or historical note;
> Preliminary list of authorities; and
> Any existing information from the repositories regarding the collection or the creator, such as accession files and donor files.

[Slide 20] At the same time, it seemed like an opportunity to take this new challenge and use it to learn a little more about minimal processing and what it is about processing that really takes the time.
We created a worksheet to track what does and does not work in minimal processing. Specifically, we want to know what tasks take the bulk of time during processing, how the type and age of material affects minimal processing, if the collection was a good candidate for minimal processing, and what the students would have like to have done with the collection if they had had additional time.

We were also concerned that the repositories might be left in the dark about the condition of the collection since they would not be actively involved with the processing. Therefore, we created a worksheet to indicate the need for preservation work in each collection we process. This is given to the repository at the end of processing. In the same vein, as we began uncovering these fabulous collections, we saw that there was huge potential for material to exhibited and/or digitized. But again, we worried that the repository staff may remain unaware of these gems because their hands were not really in the collections and therefore, we created a worksheet to address these highlights in the collection.

Finally, Courtney, the project archivist, who had been involved in the survey project, thought that re-evaluating the research value of the collection post-processing would be fascinating so we now ask our students to do this as well.

All this information is being tracked in an enormous Excel spreadsheet and we have every intention of sharing the data at the end of the project.

[Slide 21] Not long ago, Courtney and I both read Rob Cox's "Maximal Processing, or Archivist on a Pale Horse," and since then we have talked a lot about our work in relation to the article. There is no question that we are minimally processing collections, but we are looking, now at our work as leaning towards maximal:

Because the processing plans and the biographical or historical notes are created before the student begin their work, our students have 2 hours per linear foot for the actual physical processing and description of the container lists. They are also responsible for writing an abstract and a scope and content note as well as enhancing the provided biographical note whenever possible. In this way, we are doing out absolute best not to skimp on description.

We also hope that the processing plan gives our students the opportunity to make the best possible processing decisions—we encourage them to discuss our suggestions with each other, repository staff, and Courtney and myself in order to ensure that the collections receive the most complete and best treatment possible in the time frame given.

Finally, Courtney and I have always seen our work as a first step—not the final product. We hope repository staff isolates collections or even segments of collections for fuller processing now that they know what is in each collection and as researchers provide evidence of what they see as the most valuable components. This is, in fact, already happening, which makes me very happy. At the same time, we are very aware of the reality of the economy and the staffing situations at many repositories not only in Philadelphia, but across the country and are therefore, trying to make certain that every finding aid we create can stand on its own merits in case there is never future work.

[Slide 22] A centralized training of project staff had also been suggested by PACSCL before I arrived … it is probably the most important standardization that we could have implemented. Our students receive training in an intensive 3 day boot camp. The first

day is in the classroom with discussions on minimal processing generally, and minimal processing specifically in relation to our project. The second and third day, we work at a repository, and process, as a group, in the same fashion that we want our students to process on their own. As a result, all of our students are trained with the same exact information and they are all aware of the expectations we demand. Moreover, this training is not repository-specific and they are able to move from one repository to another regardless of changes in size, technology and subject content of collections.

[Slide 23] This project is consortial on so many levels. I have never had the privilege of working with so many dedicated and excited people with the same goal, nor do I expect to again. The great thing about working together in this project is that the Project has benefitted enormously from the knowledge, experience and concerns of PACSCL, the project team and the repositories. I am fairly confident that we have provided benefits to them as well.

[Slide 24] First it should be mentioned that this project could not be completed without the consortial help of PACSCL … so we may be benefiting them in the long run, but first we were benefited BY them. We work essentially with three committees: the Oversight Committee who addresses our largest concerns and makes some of the really hard decisions for the larger PACSCL group; the Standards Committee who met with me regularly right after I started the project and helped me with manuals and decisions that needed to be accepted by all 23 participating repositories; and the Technology Committee, a group of tech savvy PACSCALians who are on stand-by for when technology gets the better of Courtney and me.

[Slides 25 and 26] I would like to take just a few moments to talk about our team—because none of this could be accomplished without their hard work, dedication and enthusiasm. Our project archivist, who is the only other full-time position on the project is Courtney Smerz. We are on our second team of processors. Our first group was hired in October of 2009 and 3 of the 4 have just graduated with their MLIS degrees. Our second team, who started on June 1, are currently working at the Free Library of Philadelphia and the College of Physicians of Philadelphia. We also have an Archivists' Toolkit cataloger whose job it is to enter legacy finding aids into the Archivists' Toolkit.

[Slides 27 and 28]

[Slide 29] Working with the amazing repositories in this consortium has been an honor … I would like very quickly to thank Haverford College, Drexel University, and Drexel University College of Medicine who volunteered to be our "guinea pigs." We started processing their collections last October and they truly had faith in the project!

[Slide 30] In return, the repositories get a lot of lessons learned about minimal processing and what they might realistically consider doing in their own repositories … and probably more importantly, what they should NOT consider doing in their own repositories.

[Slide 31]We have processed x number of collections at x number of repositories [slide]

[Slide 32] We are well on the way to creating an EAD repository—which is really one of the most tangible products of the project.   The University of Pennsylvania's DLA has created this site and I am extremely excited about it. I have one small story which I would like to relate to you and showed me how valuable a resource this will be. Tell about the ANSP collection [ornithologist with WWII drawings]. Researcher of all types will no doubt find this site to be a great starting place for their primary source research! This should, fingers crossed, go live in early August!

[Slide 33] Implementing the Archivists' Toolkit. (slide text). There has already been some talk of a PACSCL-AT support group—hopefully alleviating concerns which I raised earlier in this talk about adding new software to repositories without IT staff.

[Slide 34] We are really hoping that this project makes repositories aware, not only of what they have, but of what others have too … the interconnections of the historic Philadelphia folk is pretty amazing.

[Slide 35] I have talked a lot about how PACSCL is benefiting from our work, but I am sure you are all wondering how you might benefit … we hope that much of what we have done can be adapted by other repositories, large and small, with or without technology, and with or without extensive staff.

Primarily, the most exciting thing we learned is that minimal processing does work for pre-20th century manuscript collections, and can be applied to collections of almost any age. More important than age, we have found, is the original order of the materials. We are developing a list of questions that an archivist can ask to determine if a collection is or is not a good candidate for minimal processing, which is available on our website. Courtney and I both feel that you can accomplish more with minimal processing than either of us expected—the finding aids can be fairly thorough and researchers will have a much greater chance of using materials that have been minimally processed than those that have not been processed at all.

[Slide 36]As mentioned before, the centralized staff and workflow has been crucial for the success of this project, and I anticipate, it could be very useful for repositories with large volunteer staff, or even full/part time repository staff. A brief outline of our training is online for others to see.

Very intentionally, all of Courtney and my documentation is done in Microsoft Office. Our processing plans are created in Word and we have all our tracking mechanisms in Excel spreadsheets ... technical know-how and experience and fancy software are not necessary. Examples of all of worksheets are available online.

Our Minimal Processing Manual, in which we outline guidelines for this project, but which we believe can be adapted to any repository regardless of technology, funding and size, is available on line, as is the Archivists' Toolkit manual. I have also provided my

instructions for downloading and installing the delightfully FREE MySQL and the Archivists' Toolkit!

Obviously very few people are lucky enough to have Penn's DLA create an EAD repository, but AT creates simple html pages and pdf documents right out of the box--if you have a web page, you can put finding aids on the web.

[Slide 37] Don't want to minimally process collections at 2 hours per linear foot? I don't blame you and I don't recommend it ... I tend to think of what we are doing as Extreme Processing where we are testing the limits of what can be done. However, it is important to state, that if we did not have the time and money constraints of this project, we would not be working at this pace.

[Slide 38] in fact, Courtney and I don't recommend placing a time on processing at all--rather, define a level of processing that you want completed for an individual collection--what will best benefit your researchers? Do they need item level description? If yes, then that is probably what you should do. More likely than not, good folder titles will probably get your researchers to the right place. Determine what sorts of preservation is NECESSARY. Do you have temperature controlled environment? If you do, your paper clips will not rust anytime soon, and possibly a few steps can be skipped so that access to the collection is achieved sooner.

[Slide 39] Probably the best way to learn from our mistakes, as well as our successes, is to keep an eye on our web page. One of the things that I was most adamant about in this project was honestly describing our work. I have encouraged everyone with whom we work to write blogs and I post them … they do not need to be flattering and they might criticize our work, but I know for a fact, that what does not work is often as important as what does work.   As I have mentioned earlier, almost all of our documentation is available on our website for others to use, adapt, or simply peruse.

[Slide 40] And now, at long last, I arrive at the reason we are doing all this work … for the scholars and students who will, hopefully very soon, be using these amazing collections. Until the finding aid site is up and running (hopefully August), our website is our main outreach tool. We have worked really hard to encourage our processors to write blog posts for the collections on which they work. At this point, it is really the only way that we have to "unhide" the collection and provide some hint of the content contained within a collection.

We have also created a page where we have organized the collections by topic, so that researchers will hopefully start to see all the material available on their chosen topic.

Courtney and I are also active judges for National History Day and are working with the NHDPhilly coordinators in order to make certain that students are made aware of these new resources for their research.

We have a million ideas of how to reach out to students and feel that once the finding aid site is live, we will be able to more aggressively promote the collections that are being revealed via this project.