How I Learned to Stop Worrying and Love the Backlog: Using Minimum Standards and Appraisal to Process Modern Collections

How I Learned to Stop Worrying and Love the Backlog: Using Minimum Standards and Appraisal to Process Modern Collections

Tom Hyry, Manuscripts and Archives, Yale University Library
Society of American Archivists 2004 Annual Meeting
August 6, 2004

The title of my paper today is "How I learned to stop worrying and love the backlog." Like many titles, it is meant to be provocative and not necessarily to be taken literally; I do not love my backlog, but like many long-term committed relationships, we have learned to live with one another. By using this title, I do want to get at a question of perception. Our backlogs are caused by many factors, often beyond our control. However, a part of our problems lie in the way we understand our backlogs and processing standards in general. I hope to challenge these conceptions through describing an approach to processing and backlog management and perhaps more importantly, to understand this approach within the context of evolving processing standards. A part of my argument today must be understood as semantic. Part of the way we will overcome the backlog problem is to come to different conceptions of the terms "processing" and "backlog." My paper, then, will cover the following four areas: 1) consider some of the forces causing our current backlogs; 2) suggest some principles archivists can use when approaching their backlog; 3) take a brief look at some past and current archival writing on the topic of different levels of processing; and 4) describe the "minimum standards" approach adopted by my repository

In my introduction to this session, I cited some survey results that show the depth of the backlog problem. There are several forces causing our backlog problems and I would like to offer the following list of six truisms to articulate these issues:

1) Collections are getting bigger. Although the concept of the paperless office has been with us for many years, anyone who works with modern records knows that computers create more paper than they save, to say nothing of the impact of the photocopier and other means of easy duplication and records creation.

2) Collections are getting more complex, as a variety of media, both electronic and analog, are now routinely included among collections.

3) Most repositories still collect more than they process each year. I don't have any hard data to back this up but it is the nature of research libraries to have our collecting reach exceed the grasp of our technical services capabilities.

4) Archivists are being asked to do more than we ever have before. Here I am thinking of the additions of digitization services, expanded outreach and instruction activities, and marking-up of finding aids to our duties, to name just a few of many new tasks.

5) Budgets, for the most part, are not getting bigger and at least are not keeping pace with the growth of collections and duties.

6) In this circumstance, something has to give. What that usually means is processing because it occurs behind the scenes and is often less pressing than other functions on a daily basis.

These forces, when added to the fact that most of us have never had an adequate amount of staffing to keep up with processing new accessions, have resulted in a serious backlog problem for the profession, evidenced by the intense interest shown in the ARL working conference that Barbara described. One way to deal with these realities must include adapting our processing methods accordingly. Before describing a specific approach, I want to offer the following principles and general suggestions, which can serve as a basis for managing a backlog, particularly one filled with large modern archival collections.

1) Every collection in a repository deserves some level of description. Following that assumption, it is more important to have at least a minimal level of processing and description done on all of a repositories holdings than it is to have highly detailed work done on a few collections, while others suffer the fate of being hidden from researchers.

2) A corollary to this principle is that not all collections are created equal. In a perfect world, we would process all of our holdings to an ideal level. In the real world, we must make hard decisions about which collections to process more fully, and which can get by with less. These decisions should be closely linked with collection development policies, repository mission statements, and other appraisal tools and methodologies that should be thought-out and articulated in advance. Our most important and difficult collections still need detailed processing, with the realization that these collections will make up a small part of the whole of our holdings.

3) Similarly, one does not need to process all parts of a collection to the same level. Even collections with extremely high research value could include series within them that do not need detailed work.

4) My next point is that collections are not truly available to researchers unless they are described on-line, at least at the collection level, through catalog records, finding aids, or other tools. Repositories who lack the resources or inclination to do MARC cataloging or EAD finding aids should still make their finding aids

available on the web, where they can at least be searched through Google or other web search engines.

5) An obvious statement: we process materials so that researchers can use them. Our decisions about processing should be driven by how well a collection can be used and how much it will be used.

6) Finally, we should make our preliminary inventories, even if they are only sketchy accession records, available to researchers, ideally with the same mechanisms we use to make the finding aids for our processed collections available. To do this, we may need to sacrifice some of our pride in craft. From a user perspective, imperfect information is better than none at all. Moreover, we should allow researchers access to our unprocessed and under-processed collections, unless there is a very compelling reason not to. Such reasons include collections with materials that are restricted, either by donor agreements or by law, collections that have very severe preservation problems, or collections with major potential for theft. While there are exceptions, our default should be to make collections available, whether they have been processed or not.

These are not new ideas. A friend told me this maxim: if you think you have a new idea, it probably just means you haven't read enough. The idea our processing methods and expectations need to evolve has been with us for at least a generation. Others have written persuasively on the topic and many of the principles I just articulated echo earlier work. In a 1982 *American Archivist* article, Helen Samuels and Karen Lynch describe an "adaptable approach" to arrangement and description that stresses a continuum ranging from item level to collection level processing. In the model they developed at M.I.T., the amount and type of processing work for each collection varied, depending primarily on the research potential of a collection, with the overall pragmatic goal of making the collection usable for researchers.[1]

Despite the fact that this approach was described in the heavily-circulated M.I.T. processing manual and cited favorably by Frederic Miller in his basic SAA manual on arrangement and description[2], evidence suggests that these recommendations have not been adopted as widely as they should. A recent study conducted by Mark Greene and Dennis Meissner about acceptable processing rates reveals that even when faced with immense backlogs of unprocessed collections, large percentages of archives routinely include item-level appraisal, arrangement, and conservation procedures within their standard practices.[3]

Greene and Meissner's research deserves much more time than I can devote to it here, so I would encourage anyone interested in this topic to also attend the session featuring the NHPRC Research Fellows tomorrow afternoon. In addition to the survey report, they also make recommendations for a new set of processing standards designed to make a serious dent into our unprocessed collections, what they refer to as the "golden minimum." This set of practices avoids much of the time consuming work of processing, such as arrangement within folders, re-foldering, preservation photocopying, fastener removal, and other item level conservation techniques.

Greene and Meissner's piece reads as a call to arms, suggesting that archivists need to abandon procedures we have traditionally used to process

---

[1] Slotkin, Helen and Karen T. Lynch. "An Analysis of Processing Procedures: The Adaptable Approach," *American Archivist*, Vol. 45 (2), Spring 1982, pp. 155-163.
[2] Miller, Frederic M. *Arranging and Describing Archives and Manuscripts* (Chicago: Society of American Archivists), 1990.

collections. Their work also dovetails quite nicely with the ARL Hidden Collections initiative that Barbara described earlier.[4] One of the central urgings that came out of the Hidden Collections meeting last September was the need to share local models, and I would like to devote the rest of this paper to describing an approach to processing and backlog management adopted by my shop, Manuscripts and Archives in the Yale University Library.

About five years ago, my colleagues Christine Weideman, Diane Kaplan, and I were charged with coming up with a way to handle a backlog of unprocessed manuscript collections that had accumulated over the past few decades. The backlog consisted predominately of 20th century collections, ranging from 20 to 200 feet and beyond. Unprocessed at the time for us meant collections for which we had a collection level catalog record and at least some inventory, either at the box or folder level, or somewhere in between. No preservation, arrangement, or appraisal work had been done, and the collections only had rudimentary description through what we call a "preliminary" finding aid. As we began our analysis, we were working on the assumption that each of these collections would receive it's day in the sun, that we would one day get to them and give them a full processing treatment, even though many of them had been in the backlog for twenty years or more without any demand for use or further processing.

Within this project, we took a quick look at each of these collections and quickly determined that there was no way that we could possibly give full processing to each one. Using even a high rate of processing speed to make estimates, we determined we were looking at 10 to 15 years of work ahead of us,

---

[3] Greene, Mark and Dennis Meissner. "More Product, Less Process: A Low-Calorie, High-Fiber Alternative to Traditional Processing Expectations" Unpublished paper presented in draft form at the NHPRC Archival Research Fellowship Symposium, Northeastern University, Boston, MA, June 11, 2004.

and only if we stopped collecting all together. Clearly, different methods were needed.

As a response to this problem, we developed several different categories within our backlog:

Frontlog: New collections or additions that require only a modest amount of work.

Professional: Collections with an extremely high research value, or have the most complex problems in appraisal, arrangement, description, and/or preservation.

Graduate student: collections that could use more descriptive analysis or appraisal, but that are small enough to be handled by a part-time person over the course of an academic year. Some of these collections also require language or subject skills.

Finally and most importantly for this paper, we established a category that came to be called "minimum standards" for processing, which we planned to use on the majority of the collections in our backlog. These minimum standards included the following stripped down tasks.

1. Basic arrangement of the collection into a sensible whole, at most to the folder level.
2. Folder level inventories are created.
3. Only major preservation concerns are addressed.
4. Restricted materials are located and segregated, if they are easily identified.

---

[4] Information on the Hidden Collections meeting and Association of Research Libraries' initiative can be found at http://www.arl.org/collect/spcoll/ehc/info.html.

5. A finding aid is written that includes an overview of the papers (which we use to construct a MARC record) and a box and folder level inventory.

This level of processing does not include much appraisal, unless it can be done on the series level. We also do not meticulously remove metal fasteners, arrange items within folders, or write long descriptive scope and content notes or biographical sketches. The philosophy is to streamline the work, in order to do only what is absolutely necessary in order to make our holdings more usable as quickly as possible. While a professional archivist develops the processing plans for minimum standards collections, most of the work is overseen by a support staff member who in turn supervises students to carry out most of the work.

We had initially planned to attack the minimum standards backlog aggressively, but like many great plans, circumstances have changed our conception of the backlog even further. We successfully processed a few of the collections in the minimum standards backlog, including a fifty foot collection relating to a former faculty member. However, our use of minimum standards has had a much greater impact on our newly accessioned collections than those in the backlog. This has occurred for at least a few reasons. First of all, because many of the collections in the minimum standards backlog were those which would not meet the more stringent appraisal guidelines we use today, we feel comfortable leaving them in their current state of rudimentary physical and intellectual control, until and unless we have great reason to do more work on them. Even though some of our preliminary inventories leave much to be desired, these collections *are* available to users. When we do decide to process these collections to a greater level, we base those decisions primarily and almost exclusively on the perceived research value and use of the collection. We no longer assume we will give even minimum standards treatment to all of our collections and we have current plans to further re-evaluate these collections to determine their research value and prioritize them accordingly.

Finally, and probably most importantly, all of the staff members assigned to work on processing the minimum standards backlog (and the professional backlog, for that matter), have been pulled away from processing by other more pressing duties. Chief among these are our efforts to convert the legacy finding aids for our ca. 2400 collections into EAD, so that we can make them available to the research community via the Web. These finding aids include both our fully processed collections and our preliminary inventories. While we would ideally like to do more work on certain collections in the backlog, we feel comfortable prioritizing other work over backlog processing because we do have basic control over and provide access to these collections.

Despite the fact that we haven't done an ideal amount of work on our minimum standards backlog, we have used the minimum standards concept to improve our accessioning procedures. We now take pains to do more work on collections as they come into the repository. It has been our policy for years to create basic catalog records and at least box-level inventories for all newly accessioned materials. We now strive to provide more detail in our preliminary finding aids and catalog records, fully conscious of the fact that we will not get back to many of the collections we accession. We also have used the minimum standards concept when communicating with donors, to educate them about the realistic level of work we will do on their papers. Not only has this helped us to avoid angry donors who feel we haven't worked hard enough on their records, it has allowed us to more easily use donors to help us do some of the work, when possible. On several occasions, we have been able to use donors to re-house collections, provide simple arrangement, and create inventories. On these occasions, our work gets reduced to editing, quality control, box labeling and bar coding for storage in our shelving facility and the incorporation of the inventory files into our finding aid system. Colleagues who work primarily with university archives holdings have had even more success with these methods, as they now require university offices to re-box materials and create folder level inventories for all accessions, as a condition of the archives agreeing to take their records and provide services.

There are certainly drawbacks to using these methods a few of which bear mentioning. First of all, archives 101 tells us of the inverse relationship between description and reference. The less description, the greater burden is put on public services and researchers to find things. Connected with this, less detailed description requires us to do more retrieval of boxes. Another drawback relates to security. It is undeniable that having less control over items in collections makes it easier for thefts to go undetected. Related to this, by focusing on the general rather than the specific in collections, we run the risk of losing what might be called the "diamonds in the rough." We have all worked on collections that have surprised us by revealing fascinating and vital records that we did not expect to encounter. Finally, looking forward, collections processed only minimally provide a greater challenge to digitize. It's one of the ironies of our time that as the mass of information we are confronting requires us to use less detailed means of description, effective metadata for digitization projects requires greater specificity.

Despite the drawbacks, though, we are willing to live with these issues. We can always choose to process a collection more fully at a later date, should a powerful need to do so arise.

Stepping back again, I would like to articulate a few other areas of consideration of moving our processing towards more minimum standards. A concern I have has to do with the ways we teach processing to new archivists and students. As a supervisor to a number of interns, I am painfully aware that I am still teaching students to process at a greater level of detail, constructing meticulously written scope and content notes, series descriptions and biographical statements, as well as re-foldering and item level preservation techniques. This is how I, and I would guess most of you in this room, learned to process. By continuing these methods of teaching, I fear we will reinforce processing methods that are outdated at a time when we face the immensity of modern records. However, by teaching a more minimal approach, such as I have described here, we run the risk of losing a connection to "the stuff," a connection that invariably draws new archivists into the field. I believe this question about how we conduct internships and other practical educational techniques remains open and deserves further work.

My second broader implication is more positive. The volume of modern records and the backlogs they create suggest that the largely theoretical work examining archival authority systems such as Encoded Archival Context shows great promise.[5] If we do less examination and description of actual records during processing, it will behoove our users and us if we can communicate the contexts in which records were created, to better enable them to make educated guesses about the location of records relevant to their research. I also believe that the idea of creating annotation systems to on-line finding aids gives us the potential to capture a greater level of description, which can be added by reference archivists and researchers at the time they encounter records.[6]

---

[5] see Evans, Max. "Authority Control: An Alternative to the Record Group Concept," *American Archivist* 49/3 (Summer 1986): 249-261; Bearman, David and Richard H. "The Power of the Principle of Provenance," *Archivaria* 21 (Winter 1985-86): 14-27; and the web site for the Encoded Archival Context working group: http://www.library.yale.edu/eac/

[6] see Light, Michelle and Tom Hyry. "Colophons and Annotations: New Directions for the Finding Aid," *American Archivist* 65/2 (Fall/Winter 2002), pp. 216-230.

Unfortunately, the development and management of contextual and annotation systems adds yet another project to our already overburdened agenda.

Finally, I would like to make a clarification of something implied earlier. I do not think that full processing of collections is dead, far from it. We must continue to provide a high level of work to our most important collections. However, we need to be very selective and intentional in deciding which collections get this work and those decisions should be both tied in with each repository's appraisal guidelines and collecting policies and also documented in repository policy and perhaps in our finding aids themselves.

To return to the introduction of my paper and its title, a sensible conceptual approach to a backlog that includes varied levels of processing and the development of reasonable expectations that we can communicate to users and staff lifts the burden of backlogs considerably. While no one will ever actually love a backlog, if we can succeed at getting a basic level of control over our holdings and make them available to users, it frees us up to work on the richest of our collections. To push this idea a bit further, using the term backlog to describe these collections both misrepresents our realities to outsiders, thereby raising expectations unreasonably. It also discourages archivists, who feel we are always running to catch up or worse, failing. Perhaps the best way to make backlogs disappear is to stop calling them backlogs in the first place.