Organizing Large Contemporary Manuscript and Archival Collections: Ideas for Planning Work

John Dojka
Yale University

Follow this and additional works at: http://digitalcommons.kennesaw.edu/georgia_archive

Part of the Archival Science Commons

Recommended Citation
Available at: http://digitalcommons.kennesaw.edu/georgia_archive/vol10/iss2/4

This Article is brought to you for free and open access by DigitalCommons@Kennesaw State University. It has been accepted for inclusion in Georgia Archive by an authorized administrator of DigitalCommons@Kennesaw State University. For more information, please contact digitalcommons@kennesaw.edu.
The challenges posed by the volume of contemporary documentation have become familiar to all archivists. The district congresswoman, retiring next year, seeks a repository for the four or five tons of records documenting her thirty years in Washington; a once important local industry closes its doors, leaving five or six thousand feet of production, accounting, and personal records mouldering in a leaky warehouse; a local social service agency is moving to new quarters and wonders whether a research institution would be interested in taking the several hundred feet of case files stored in its basement.

The decision to accession all or even part of such a body of records will have serious consequences for the future of a repository, especially in terms of budget, staffing, and collection policy. Yet, the alternative is to stand by and watch an important part of a town's or region's past be turned into landfill or sent off to the paper recycling plant.

A repository's response to such a situation should be determined by collection policy, stack space, budget and staffing priorities. Another factor, the significance of which is frequently unacknowledged or unperceived, is the way in which a repository has organized its processing operations.

If processing procedures are organized on a model geared to arranging and describing records at the document level, an archivist might reject automatically large bodies of contemporary documentation. Such collections would pose a major threat to a repository, placing unbearable demands on its resources. On the other hand, if processing operations are flexible so that
they can be used imaginatively with a variety of documentation, the challenges posed by contemporary records will not appear insurmountable. Collection policy and other priorities, not processing operations, should be the determining factor in accessioning decisions.

What follows, in the form of several premises and a brief work plan or model, are some ideas on how processing operations can be organized to deal effectively with the problems posed by the volume of contemporary documentation. The model can be applied to personal papers and archival records as well as to organizational records whose archival integrity has been destroyed. It can be used by archivists and manuscript curators in a variety of settings.

The work model is both a planning and an implementation tool. It provides guidelines for projecting and planning the priorities and flow of work. It offers a framework for breaking the work into its component parts to estimate staffing, budget, supply, and time requirements as well as to allow easier supervision of staff and assessment of work progress. Furthermore, if assessment of existing resources indicates that one must look outside the parent institution for funds and staff, the work plan will help to demonstrate to a funding agency one's capability for organizing and successfully carrying to conclusion a large processing project. In short, it is a means of conceptualizing processing operations that enables one to see the potential order in unprocessed collections and to plan how that order will emerge. It is not meant to be a detailed discussion of specific processing mechanics. Those procedures and the necessary supplies have been well described by Kane and Duckett.

The premises underlying the model are, first, that although manuscript collections are the unique creations of a unique individual or institution, standard processing procedures can be applied to collections without destroying their integrity. The argument that collections are unique has too often kept archivists from focusing on common elements and planning processing operations accordingly.

Second, at the outset the archivist must have a
clear view of the goal—the level of control and description it is feasible to achieve in a reasonable period of time with limited resources. It is not necessary to read or even handle each document in a collection to obtain adequate intellectual and physical control of the materials. The size of contemporary collections mandates a turning away from a document level mentality of processing to a view of processing as part of a repository's total mission to collect, preserve, and make available a representative record of human experience. There is a sort of equation here—the more a repository uses its budget and staff time to process to the document level, the less of its total resources it is able to spend for collecting and making other materials available to researchers. A conscious balance must be struck between each of these priorities; the processing of contemporary collections must be viewed as inherently a compromise. This may mean control and description at folder or even series level for an entire collection.

Even within the same collection different series may warrant greatly varying levels of description. For example, in a body of congressional papers the personal correspondence of the legislator may well warrant folder or even document level description while a series composed of several hundred feet of case files or public opinion mail might be adequately described in a paragraph or two.

The danger is, of course, that when approached with a document level mentality, one 300-400 foot collection (by no means uncommon for a contemporary figure) poses a major, even crippling, commitment for a repository. In the end, that repository will have either a huge backlog of material, unprocessed and unavailable to researchers, or it will curtail its collecting scope to exclude large contemporary collections. For most repositories, neither of these alternatives should be acceptable. A well-designed processing operation can head off such problems.

Third, reviewing records for potential weeding and sampling projects should be incorporated into processing operations as a routine step. Fear of using
weeding and sampling techniques often seems based on a fundamental misconception of the role the archivist plays in the production of historical knowledge. On a workaday level archivists are inclined to forget that history is a human creation based on a selection of events woven into an operational explanation of the past. The archivist's goal, as Gerald Ham has stated, is "an informed selection of information that will provide the future with a representative record of human experience in our time." ²

The most significant aspect of the archivist's work is that through collection, preservation, and reference policies one selects evidence from which the researcher in turn makes a selection. Weeding and sampling projects, imaginatively conceived and proficiently carried out (perhaps with the aid of historians and statisticians), are by no means inconsistent with this goal. In fact, a good deal of contemporary documentation may not even be usable by researchers prior to a sampling project that reduces it to manageable proportions.

Fourth, because a major portion of processing large contemporary collections is very routine, work should be planned and structured to allow delegation of as many repetitive tasks as possible to part-time assistants. This does not mean that the archivist's own hands do not get dirty or that the job becomes merely supervisory. Rather, it implies an awareness of which decisions and procedures call for professional expertise and which may be handled more efficiently by someone with less training.

The processing operation should be seen as a series of steps, the aim of which is to gain successively more sophisticated, intellectual, and physical control over the material being organized. The implementation of the work plan is the unfolding, in stages, of the inherent order in an unorganized body of records. There are six major phases of work:

1. Background research
2. Inventory of records and preliminary grouping
3. Identification of series and arrangement on paper
4. Review for weeding and sampling
5. Physical arrangement and processing mechanics
6. Preparation of the finding aid

Background research does not mean an exhaustive investigation. Instead, the aim is to construct a brief, chronological outline of the pertinent points in a person's career or an organization's development. In the case of an organization or institution an effort to construct an organizational chart may be worthwhile. Gathering the data may be as simple as consulting a volume of Who's Who, or it may require a search of more comprehensive reference tools.

What this outline of information provides is a potential map of the collection. It alerts the archivist to the type of materials one may reasonably expect to find, gives an explanation of materials that might otherwise seem inexplicable, and provides information for later arranging the collection into meaningful series.

After acquiring experience in dealing with a variety of collection types, given knowledge of the person's career, the type of business or institution, one should be able to predict the type of materials that will be encountered and the potential series.

Moving through the next phases of processing, the archivist fills in the biographical or organizational outline with relevant information. These data will form the nucleus of the biographical or historical sketch that will introduce the finding aid.

The next priority is to produce an inventory that will provide an overview of the records. The archivist may be confronted with one of three situations: (a) the records or personal papers will be found with overall order intact—i.e., filing units, groups of correspondence, photographs, financial records preserved; (b) the records or personal papers will be in such disarray that any once-existing order is not apparent; or (c) a combination of a and b.

When the original order of the records or personal papers is intact, conducting an inventory is a rather straightforward task of moving through the records, boxing them if necessary, and listing the following information for each box, file unit, or group of material: type (correspondence, incoming; correspondence,
outgoing, copies; printed material; subject files; research notes; financial records; organization of the records (portion of an alphabetical subject file, chronological); broad subject area of the material; inclusive date span; and estimate of work to be performed (preservation, photocopying, refoldering, alphabetizing).

When the records are in complete disarray, a preliminary sorting will be necessary before likely file units or series appear. Essentially, the archivist must impose an order on the records by rough sorting of all materials into categories and boxing or reboxing them according to groups. Many contemporary collections are foldered, with captions, making an inventory fairly easy. However, folders should be spot-checked to ensure accurate captions. If materials are loose and unfoldered, sample handfuls to obtain the necessary information.

It is imperative to resist the temptation to do extensive sorting at this point; merely block out the main types of material, sorting into major groups. In the case of institutional or business records, the major groups will be minutes, constitutions, voting records, correspondence, project files, financial and production records. With personal papers, likely groups will be correspondence, photographs, financial records, subject or research files, and memorabilia. Care must be taken to avoid breaking up units which have an obvious relationship so that they may be preserved for more thorough inspection. Another useful procedure is to physically lay out the records, looking for similar filing tabs that represent a once-existing filing scheme. It is important, however, not to get bogged down trying to reestablish filing schemes.

The archivist should also be wary of another danger present at this stage—the temptation to begin minute sorting within the major groups. Most major groups can obviously be broken into finer units; for example, the financial records of an industrial institution can be divided into accounts receivable, invoices, payroll. Don't give in to the temptation! Remember that the goal at this point in the work is to establish a broad, primary control over the records. Reserve more detailed
sorting for later. Once the major sorting is completed and records are boxed, information on type, organization, subject area, inclusive date span, and processing mechanics should be listed for future reference.

Finally, the archivist may confront records or papers in which portions of the collection have their original order intact and other portions are in disarray. In this case, sort out and preserve the intact units and, then, apply rough sorting procedures. In all three situations, the goal is the same--a list of broad groupings or units from which series can be constructed.

Once the inventory is completed, materials in the collection can be manipulated on paper to form a preliminary arrangement by series. The purpose of this step is to refine the groups identified in the initial inventory, breaking them down into series and sections of series to establish finer control. This is essentially a process of examining the inventory for patterns and either grouping material that can be related on the basis of subject or form or deciding to preserve the original order of the records intact.

In the case of an organization's records, for example, correspondence can frequently be divided into series of incoming and outgoing according to discrete alphabetical runs covering different date spans. Policy records can be broken into series of annual reports, committee minutes, planning reports; financial records into series according to form (daybooks, ledgers, audit reports); organizational publications and printed material into chronological runs. Once this further refinement has been accomplished, review the internal organization of each series. By referring to inventory notes, make lists of the type of conservation and processing mechanics that must be performed on each section--refolding, breaking boxes of loose correspondence into chronological runs, alphabetizing runs of subject files. On the basis of the data gathered, review any earlier decision about the level to which records will be organized and described: Does the material warrant description at the folder or the series level? Hard and fast rules cannot be established. The level of description will depend on the resources.
and time available and on the condition of the records. In regard to organization, imagination should be given free rein to consider the elimination of as many work procedures as possible. Do staples, unless rusty, have to be removed; if folders are in good condition, are acid-free ones really necessary; is it imperative that materials be arranged chronologically in folders? If certain series on low quality paper are to be filmed for conservation, can staples be cut off in large batches with a paper cutter instead of being removed one by one?

The next step is to appraise each series for possible reduction of bulk by weeding and sampling. By weeding, can duplicates, envelopes, memorabilia, printed matter, and photographs be removed for processing in another section of the repository or for destruction? Does the amount of bulk that will be reduced by removing ephemera warrant the time involved?

The purpose of this paper is not to discuss specific sampling techniques in detail. Rather, it is to point out when sampling should be considered to reduce the bulk of a series, what the basic types of sampling are, and what problems may be involved. The two general types of sampling that can be applied to records series are qualitative (selective) and quantitative (statistical) sampling. It is vital to recognize the characteristics and limitations of each method. There is no such thing as a sample of general utility that will satisfy the needs of all prospective researchers. The sampling method chosen will limit and determine how the material sampled can be used for research purposes.

The possibility of sampling will arise when a large series of records, similar in form, is present--case files, forms, correspondence, financial records. The question is basically one of appraisal: Are the contents of each folder unique, and does that uniqueness have significant historical value? If the answer is yes, sampling should be ruled out.

The sampling method used will depend on anticipation of the potential use of the material and its degree of homogeneity. If, for example, the contents of the series are homogeneous and the aggregate of the information contained in the records, not the individual record
itself, will be of importance to researchers, quantitative (statistical) sampling methods can be applied to obtain a sample that will reflect the characteristics of the whole group. Two methods exist for obtaining this type of sample: Systematic (choosing samples in a predetermined numerical sequence, say every tenth, fifteenth, or twentieth file) and true random sampling (using a table of random numbers to insure efficient distribution).

Systematic sampling, if administered correctly on a large series, will yield a sample that is essentially random and representative of the characteristics of the whole. However, a potential problem with its use arises in the case of files arranged in some numerically repetitive pattern or cycle, so that perhaps every twentieth file is somehow similar throughout the whole series of records. Potential bias can be avoided by using true random sampling with numbers selected from a table of random numbers, such as the Rand Corporation's A Million Random Digits. Properly administered, this will guarantee a highly reliable sample and is the preferred method.

On the other hand, if appraisal indicates that the potential value in a series is specific material, not the information in aggregate, qualitative (selective) sampling may be used as a means of separating desired information from the series. Qualitative sampling involves selecting material to be preserved on the basis of some predetermined criteria of significance or atypicality—economic status, ethnic group, geographic distribution, importance of a person or group. The basis of possible selection is almost limitless, and very sophisticated strategies can be developed.

The problem with this type of sampling is that it is in effect a calculated risk or gamble—in many cases justifiable—that the only research potentials the records have are those that the sample is designed to select and preserve. Experience has shown that the research potential of records is notoriously difficult to predict. Needless to say, a decision to selectively sample records demands a thoughtful and imaginative appraisal with the realization that certain research potentials will be destroyed, among them the general quantitative use of the
In large contemporary collections sampling can dramatically reduce the bulk of materials that have been appraised at low value, such as routine requests for information, letters of congratulation, Christmas cards, certain types of form letters, invitations. In such cases a selective sample—one year's worth or one letter of the alphabet—can be obtained to show that this type of material was present in the collection or that a particular function it represents was carried out.

Once sampling mechanics have been designed and the size of the sample determined, assistants can select numbers, number folders, and do the actual pulling of the sample. Notes on methodology should be kept for the register. Sampling procedures can be very simple or complex. The archivist may feel more comfortable having a sampling project reviewed by a historian or statistician. It is one potential tool to deal with the bulk of contemporary collections that should be carefully considered.

Referring to the plan of arrangement, begin the physical arrangement of the material into the appropriate series and series sections. In some cases, this may merely involve manipulating the records boxes into appropriate order; in others, it may mean rough sorting contents of boxes to gather material for the various series. Continue to avoid the temptation to begin detailed organization of records within series and sections. The object is to group materials physically in a pattern conforming to the plan of arrangement and to renumber and shelve the cartons so that the processing work space is cleared.

One can now begin distributing series or sections of series to assistants who will perform the more routine mechanics of processing. A checklist of the procedures to be performed on each series should be prepared for assistants by revising the worksheet compiled for each series during the initial inventory. Complete any work involving decisions on arrangement prior to turning the materials over to the assistants. Tasks for assistants should be simplified to the point that they require relatively little instruction and can be performed by two or
more people working at the same time or on different schedules.

While this phase of the work is going on, spot-check for accuracy, review finished portions of the records, and take notes on subject content. Preliminary tasks on other collections can also be performed so that new material will be ready for assistants as soon as the collection they are working on is completed.

Work on large contemporary collections should be a team effort. At the center of the team is the archivist—planning, organizing, making decisions on questions of arrangement, and delegating appropriate tasks according to the abilities of the assistants. The key to successful processing in this fashion is a clear understanding of which decisions and procedures warrant the archivist's time and which can be delegated.

In many cases reserving part of the processing budget for part-time help is a more effective way of dealing with quantities of routine work than hiring another full-time staff member. The reason for this is that once a body of records has been broken into series, most of the remaining work will involve routine tasks. Such labor can be amazingly tedious, and full-time staff members cannot be expected to work at maximum efficiency without becoming bored and making time-consuming mistakes. Part-time help by conscientious students or volunteers is a more promising alternative.

To avoid the pitfall of investing too much time in supervising assistants and structuring work schedules, block out work on the project into units requiring little initial instruction and supervision so that assistants can come and go according to their own schedules. Because of the routine nature of the tasks, assistants' working time, when possible, should be limited to two or three hour periods. Naturally, assistants' abilities and levels of interest will vary a great deal, and some can eventually be given much more responsibility. However, the general guideline should be to keep the work as simple as possible, requiring the minimum amount of supervision.

Avoid taking on students under the guise of teaching them the archivist's trade. Teaching processing
properly requires a major commitment of staff time and
energy. Before a repository embarks on such a pro-
gram, priorities must be weighed. Are part-time
processing assistants to be viewed as an extension of
the processor, hired to relieve him of much of the rou-
tine processing, or as apprentice archivists? To cope
expeditiously with the volume of large collections, the
priority may well have to be moving paper, not training
archivists. Clear lines should be drawn between these
two goals.

Another major factor to consider carefully in plan-
ning effective processing operations is the use of work
areas and equipment. At large, well-planned reposi-
tories, processing areas are set up within easy access
of loading docks, freight elevators, and commodious
work spaces. Although such conveniences are in the
realm of fantasy for most archives, much can be done to
make optimum use of cramped work spaces in small re-
positories. For example, long corridors or aisles are
ideal places to string out collections for inventoring.
Equally important are such items as long collapsible
tables, sorting bins, pigeon holes, and hand sorters.
The most important factor, however, is the way equip-
ment and furniture are arranged in the processing area;
the tendency of staff to want fixed, immovable proces-
sing stations or desks should be fought. Flexibility is
the key—tables, desks, and sorting equipment should
be thought of as mobile components to be strung out to
form large, extended work areas during the initial
stages of processing and then contracted into compact
units for assistants to work at once the collection has
been broken into series and the parts not being worked
on are shelved. In short, work procedures should dic-
tate the layout of the processing work space, not vice
versa.

The archivist begins to write the finding aid when
doing the background research described in the first
phase of this work plan. Note taking on subject content
of series continues throughout all work phases. The
format of the finding aid or collection inventory will
depend largely on decisions regarding the level of des-
cription and arrangement. 6
Once work on all of the series of a collection has been completed and any final odds and ends integrated into the appropriate series, the collection can be put in order--series by series, folders numbered and stamped, and folder caption lists compiled by assistants. Assigning permanent numbers to folders and boxes must be the last step because of possible changes. Proofing by another assistant should follow each step. After a final review, the series content and physical description notes are pulled together and organized in final form for inclusion in the finding aid.

The above model is not meant to be applied dogmatically. Specific work procedures will necessarily vary from repository to repository and from collection to collection. The model is used to illustrate that large contemporary collections need not intimidate even a repository with modest resources if well-planned and integrated processing operations are developed. Above all, there must be the realization that imagination and compromise are essential.

Notes


3 For a detailed discussion of organizational procedures, see Kane, *Guide to Care and Administration* and David B. Gracy II, *Archives and Manuscripts: Arrangement and Description*, Society of American Archivists, Basic Manual Series I (Chicago, 1977) (hereafter cited as *Arrangement and Description*).


For ideas on format see models of inventories from other institutions, particularly the Library of Congress model of collection register developed by Katherine E. Brand. For further discussion of the register or inventory format, see Brand, "Developments in the Handling of Recent Manuscripts in the Library of Congress," *American Archivist* 13, 1 (January 1955). See also Gracy, *Arrangement and Description.*