Efficient Finding Aids: Developing a System for Control of Archives and Manuscripts

LYDIA LUCAS

THOSE OF US FROM REPOSITORIES that don't have automated finding aid systems tend to feel a bit defensive, probably a bit jealous, around those from repositories that do. Disparities in size, in fiscal and personnel resources, and in access to technical expertise have constituted barriers that fostered a perception of automation as beyond the means of most of us, and therefore beyond our immediate aspirations as well. The interests, concerns, and priorities of institutions so far advanced may seem irrevelant to the creation and use of the manual finding aid systems employed by the rest of us.

The last few years, however, have seen the development of cooperative projects offering an opportunity for participation by institutions of all sizes, all types, and all stages of development. The national repository guide and data base being developed by the National Historical Publications and Records Commission; the Washington Records Project that is preparing guide entries to records and manuscripts throughout that state; and the four state, Midwest State Archives Guide Project are examples, currently in progress, of projects designed around contributions from a number of repositories, relying upon information derived from the finding aids of those repositories.[1] Future years will see more such cooperative ventures, and with them will come the expectation that other repositories will eventually be able to structure and submit appropriate input data. Even those repositories whose internal procedures remain entirely manual may find it desirable and possible to prepare finding aids that can serve the needs of automated systems as well.

Besides being eyed as potential nourishment for automated systems, finding aids in today's repositories are beset from other directions. They must be adaptable to collections of increasing variance in size, complexity, types of materials, and content value. They must serve more users with a wider range of individual needs than ever before. In many reference situations, they

[1] Larry J. Hackman, Nancy Sahli, and Dennis A. Burton, "The NHPRC and a Guide to Manuscript and Archival Materials in the United States," *American Archivist* 40 (April 1977): 201–5; *Report of the Conference on Automated Guide Projects,* NASARA, St. Louis, Missouri, 19–20 July 1977; "The NHPRC Guide/Data Base: An Introduction," April 1979, information packet distributed by NHPRC; Lawrence R. Stark, "The Transition of Automation in Multi-Repository Guides," unpublished paper delivered at the annual meeting, SAA, October 1979; Max J. Evans, "The Midwest State Archives Guide Project: A Status Report," unpublished paper delivered at the Midwest Archives Conference, Spring meeting, May 1979.

must be sufficiently self-explanatory to permit a substantial degree of independent use by the research public.

With needs and demands nearly always far outstripping staff resources, practicality suggests that a single finding aid system should address simultaneously as many of these imperatives as possible. The ideal inventory or register must be efficient, in a format permitting the processor to record a maximum of usable information in a minimum of time and space. It must be easy to maintain, encouraging updates, changes, and addenda. It must enable a potential user to grasp the essence of a collection at a glance, to judge whether the collection includes items relevant to that user's research topic, to locate references to those items in the finding aid, and then to identify precisely that portion of the collection that contains them. The user would have been specifically directed to that inventory through some sort of repository-level index or similar access medium.

Consistent attainment of such an ideal may be beyond the scope of mortal man. There are, however, some general principles and practices that contribute to the ease and efficiency with which finding aids can be produced, maintained, and used. Comtemplating potential future automation might be one way to help bring these practices into focus and develop a rationale for their systematic application.

The Minnesota Historical Society's finding aids system is far from new; its core structure of individual collection inventories with access provided through a dictionary card-catalog has been in use since the 1920s. It is similar to the systems found in other repositories of comparable size, and no aspect of it is yet automated. But its most recent form was shaped quite consciously out of, among other things, a particular experience with automation. It illustrates the evolution of one repository's approach to finding aids, and the adaptations it made in both theory and practice to anticipate future automation. At the same time, we have found it to be an efficient and effective manual system.

A dozen years ago, we were preparing what might be called "scholarly" inventories. The core of each inventory was a narrative description of the collection. Often, the description was a sort of research essay in which the processor advanced through the papers, discoursing at greater or lesser length about their content, until he reached the last box and there was nothing more to be said. As a substitute for the calendaring approach, this type of inventory was informative, with extensive detail and considerable analysis and interpretation; but it had some increasingly bothersome deficiencies. The essay approach did not incorporate the concept of a separate scope note, box list, or other concise overview of the collection's character. Lacking this, it was difficult to obtain a clear sense of what the topical foci of the papers were, what materials bulked largest in quantity or importance, whether cited specifics highlighted aspects of a significant topic or represented the processor's whim, and often where they were physically located within the collection.

The card catalog, then as now based on a unit card system with fairly extensive author and subject tracings, and occasionally an analytic card for a specific item or series, was not designed to do more than direct its user to the inventory. The researcher then had to read the entire inventory, having no way to judge whether he needed to or not.

This system itself had arisen in part as reaction to a still earlier approach in which the card catalog had been used as the primary information medium. Many collections had been cataloged in more detail than they were described; names, in particular, often appeared in the card catalog without being mentioned in the corresponding collection inventory.

The finding aid system as a whole recognized only unprocessed collections, not yet entered into the system, and collections that had been described and/or cataloged in detail. For many years there was no pressing need for an intermediate status.

These problems became obvious in the late 1960s, when we became involved in the testing phase of the SPINDEX II program for automated description of archival ma-

terials. Our experimentation centered on producing a file-list printout and keyword index for forty boxes of Northern Pacific Railway, General Manager, subject files. We also formatted (but never had keypunched) some lists and indexes for selected manuscripts materials.

We bowed out of the SPINDEX project rather early. The immediate reasons were lack of funds, departure of a key staff person, and the frustrations of trying to apply and de-bug a program that was insufficiently developed for our needs. But the underlying reason was that we didn't really know what those needs were. Not only did we lack a clear conception of what we realistically could and could not expect the program to provide, we also had not clearly defined why we wanted computerization. It was surely progressive, and a Good Thing; but no gaping deficiency in our finding aid system had been defined as being both vital to our program and insoluble without automation.

In fact, we were operating under the delightful assumption that if one fed in file titles and permuted their keywords, the computer would structure the files—an attractive substitute for the lengthy tasks of arrangement, analysis, and description. Unprocessed collections could easily be brought under control, and processed collections could be indexed and correlated at will. What was missing was *context*—summaries of collections and series and types of files giving meaning to the lists and indexes that might be produced by characterizing the collection as a whole, and reflecting their relationship to it. We knew what all the file titles were, but the format offered no way to express what was in them. Nor did it provide a means of dealing effectively with lengthy series of uniform folder titles, such as "Correspondence." It was impossible to judge what a series might require in more precise file identification, indexing, and correlating, or how to devise logical control numbers. Lacking, too, was a realization that data formatting and keypunching were noticeably more time-consuming and expensive than just typing file lists, and should be expected to yield proportionate benefits. In sum, manipulating isolated file or folder-level data did not offer us a viable way to deal constructively with anything but random subject files, and, in fact, the sole survivor of our first foray into automation is the General Manager printout, which is of just such a file.

So we abandoned SPINDEX II, not without a guilty sense of being Neanderthals at heart, and began to cope with what was becoming an intimidating backlog of unprocessed collections. But in doing so, we very consciously drew on our SPINDEX experience to develop processing priorities and inventory formats that we hoped would both improve our existing manual system and accommodate whatever we might want to do in the future (including another attempt at automation, should an appropriate occasion arise).

When the occasion did arise, nearly ten years later, we were delighted by the ease with which we could prepare automated input from manual finding aids, and in turn could modify details of these finding aids to match the input format. We are currently participating in the Midwest State Archives Guide Project, an NHPRC-funded cooperative project by the state archives of Illinois, Indiana, Minnesota, and Wisconsin to produce guides to public records using the SPINDEX III program. One of the project's specific goals is to explore the practicality of developing such guides from existing finding aids of varying formats and levels of detail.

Data entry for this project is at the series rather than the file-title level. The series are placed within an overall organization scheme before being assigned the control numbers that will fix their location in the registers, guides, and checklists generated from the project's data base. The index will be cataloger-produced rather than created from keywords in existing series titles and descriptions, since another project goal is to explore the use of common terminology for certain topics and record types common to the public records of many states. The cooperating institutions record, in a standardized format, as much or as little data as their finding aids offer, and send the data to the State Historical

Society of Wisconsin for computer input and manipulation. A prototype guide from the pilot stage of this project has been produced. The implementation stage, now in progress, will enter data on as many series as possible, permitting the creation of more comprehensive guides and checklists.

I think the most important thing we have learned from these two ventures is that an automated finding aid system and a manual system are not two independent entities. They are inextricably linked. The SPINDEX III training sessions stress, and our own experience has confirmed, that a poor manual system cannot be transformed into a good automated system. In fact, an automated system should be planned as though it were going to be a manual one, exploring where and why automation will be beneficial, and how the system should be structured to facilitate it. An automated system should enhance or perhaps replace a viable manual system, not attempt to compensate for the deficiencies of an inadequate one. If anything, it must be even more simple, more straightforward, more obvious. Automation, though it tolerates wide variance in data, does not tolerate idiosyncrasy.

Our current finding aid system evolved in part out of this understanding, and in part out of the realization that the one-place narrative inventory suffered in its own way from the same deficiency as did the piecemeal file approach: lack of summary, absence of context. The system is now grounded in two interrelated principles: basic bibliographic control over all collections, and a "building-block" approach to inventories. Both principles start at the accessioning process and carry through to the final inventory for a fully processed collection.

As a collection is accessioned, a main entry including title, date span, and quantity is established; a preliminary scope and content statement is drafted; and a summary list or description of box contents is prepared. If necessary, enough time is invested in organizing the papers to ensure that the contents of each box can be defined and recorded somehow. Accuracy will be compromised at this stage, if necessary to maintain the accessioning momentum. The emphasis on basic controls reflects a paramount institutional priority: never should there be anything in our possession for which we cannot produce the information needed to identify it and its components, and to service it for the donors and the public. Also, incidentally, it helps remove a psychological burden when the backlog is no longer the unprocessed unknown but represents the first stage in a processing continuum.

Processing refines the principle embodied in the accessioning process: maximum accessibility gained as straightforwardly as possible. Inventories, especially for larger collections, are structured in progressive levels of detail, one leading into the other. Biographical or historical data and an improved scope and content note, followed by a summary box list and perhaps brief series summaries, introduce the collection and provide an overview that a user can scan quickly to judge whether it may be pertinent to the user's research. These can then be supplemented, as needed, by more detailed series descriptions, folder or file lists (selectively annotated with content notes), special lists, or citations to specific items. Description of a long run of chronologically arranged papers can be subdivided by date spans or other appropriate segments, or a general description can be followed by references to more diverse materials. Data sheets, originally intended for SPINDEX II input, now supplement some inventories as a way of recording a great deal of detail about individual folder contents, without cluttering up the main inventory.

Presenting the collection in identifiable units, and visually distinguishing the summary portions of the inventory, offer the researcher ready access to information about particular parts of the collection without burdening that researcher with irrelevant detail. Preparation of the inventory can stop, if necessary, at any level and still yield a cohesive finding aid. The archivist has considerable flexibility in selecting descriptive alternatives and depths of analysis appropriate to each collection. The format also tolerates personal predilections

and differing degrees of skill and experience among processors.

We found that conscious adoption of these two principles offered a perspective on finding aids in general. Do calendars, for instance, or lengthy folder-by-folder descriptions, provide a refuge from the need to summarize and thereby omit precious specifics? Have we been writing narratives more for their literary qualities than for functional use? Will users think us any more sophisticated because we produce either literary or endlessly detailed inventories? Scope notes, box lists, and annotations, especially for larger or more homogeneous collections, can be just as informative, simpler to prepare, and easier to scan. Use of these devices does not lessen the amount of analysis time or intellectual effort on the part of the archivist, but does offer a format whereby all of this input can be channeled toward recording information instead of being devoted to the construction of narratives or the repetition of bibliographic details.

A format that uses scope notes and series summaries to reflect the essence of a collection encourages the archivist to view the papers as a whole and to develop a concise statement of vital information that can stand independent of subsequent details. Compartmentalizing the inventories makes it possible to view them as being readily updatable, as dynamic rather than static. Revisions or additions can be made at any point and only affect a page or two. Detailed lists and appendixes, should they prove useful, can be inserted at any time without requiring that the rest of the inventory be redone. If inventories are prepared in a looseleaf format, such modification becomes a routine part of an ongoing program of finding aid maintenance and improvement. In fact, once we had become comfortable with the concepts of basic controls and levels of detail, we realized that a single finding aid system could accommodate wide disparities in information for both processed and unprocessed collections, since the system is readily updated by replacing preliminary inventories and card sets with more detailed ones.

Other considerations of inventory construction follow naturally. Inventories should be predictable, with the same basic elements having roughly the same relationships from inventory to inventory. Inventories should be easy to scan visually and refer to, with generous use of subheads, section breaks, indentations, and lists. The physical layout of inventories should help ensure that basic data are consistently recorded, by providing definite places where such data belong and where their absence will be conspicuous. Standard formats (of headings, for example), where the required elements can be formalized, help encourage precision and accuracy at crucial points. Dates and quantities should be a part of all entries, so that the reader does not have to guess.

An inventory that is seen as the core of a control system becomes the single place where a researcher can find all of the recorded information about a collection. Other types of finding aids—compiled guides and subject lists on the one hand, card catalogs or analogous repository indexes on the other—find their places in the system by linking with this core. A subject guide, for instance, should not contain information that cannot also be found in the inventory. Catalog entries should not be used in lieu of inventory references to authors and subjects.

Seeing the inventory as the system's core helps also to define the role of a repository index (in our case a card catalog), which becomes more clearly an index to all of the inventories rather than an independent descriptive medium or a means of direct access into the papers. Catalog entries are derived from information in the completed inventory, and nothing appears in the catalog that is not mentioned and placed in context in the inventory to which it refers. We have found that the ease and flexibility of preparing and maintaining a card catalog are enhanced if its use is limited to this primary function and it is not expected to offer more extensive or varied information than its format can accommodate.

A card catalog shares a particular advantage with the compartmentalized inventory: because it comes in small segments, one reference on a card, it can readily be

updated or expanded by replacing or adding cards. The cards of one collection need not look exactly like those of another, and outmoded cards need not be replaced every time a change in format is made.

Also, use of a repository index as an access medium reinforces the need to characterize each collection in a way that will clearly distinguish it from all others. For manuscripts, this is done through a distinctive title, usually combined with a catalog number. For archival records, we have established a hierarchical record group/subgroup/series/file identification that can, if necessary, be expressed alpha-numerically. Although a card catalog can accommodate lengthy collection or series titles, other types of indexes become unwieldy unless titles can be converted to symbols of some sort. This is of crucial importance when we contemplate an automation project that will require reliance on unique collection or series control numbers for manipulation of data elements.

The aim of our finding aid system is to enable us to peel off information at any point and have self-contained input for a variety of potential information systems, manual or automated, in-house or otherwise: catalogs, guides, indexes, shelflists, bibliographies, checklists, central reference services. Contemplating the requirements of automation helps reinforce this approach, for a viable automated system is structured; deals with identifiable units; requires that certain pieces of data be found in certain places, recorded in a standardized manner; does not readily accommodate long narratives; is oriented toward the discrete and specific; and takes no account of literary style or nuances.

Having developed our finding aid system this far, we now know what we need that a manual system cannot give us. These needs all go beyond what can be accommodated by more traditional finding aid formats and procedures: (1) updatability of *compiled* finding aids (as opposed to individual inventories), particularly of checklists to public records, special subject lists, microfilm sale lists, and similar consolidated reference tools that should be updated regularly; (2) effective subject access to bodies of records (railroad and public records, for example) that are hierarchically arranged and often interrelated in such a way that individual series cannot be assigned discrete names and catalog numbers; in other words, they cannot be handled readily or comprehensibly by a standard dictionary catalog; (3) subject access to random files, such as numbered subject files or related files from several series or collections; and (4) linkage with data in other institutions.

In coming full circle to a second computerization project, we have what we hope is a better grasp of our priorities, our capabilities, our needs, and the way we would like to see those needs addressed. Without a manual system that helped us pinpoint them and judge their importance, we would probably never have come to this understanding.

LYDIA LUCAS is head of technical services for the Minnesota Historical Society's Division of Archives and Manuscripts, St. Paul. This paper was first presented at the SAA annual meeting, in Chicago, on 27 September 1979.