Standards for Archival Description

From Archival Gothic to MARC Modern: Building Common Data Structures

KATHLEEN D. ROE

Abstract: Archivists have never succeeded in developing a framework that would provide consistently formatted descriptive information about archival holdings. This has resulted in confusion and inefficiency for staff members and researchers alike. The author analyzes widely used manuals of archival description, discussing the elements of information recommended for inclusion in various kinds of finding aids and the implicit "data structures" intended to contain those elements. She also reviews recent developments, mostly related to automation efforts, that have led to more explicit definitions or data structures and presents issues that must be addressed in order to to define a general archival information system standard.

About the author: *Kathleen D. Roe is associate archivist with the New York State Archives and Records Administration, where she supervises arrangement, description, and automation of the Archives' holdings. She prepared this background paper for the June 1989 meeting of the Working Group on Standards for Archival Description in response to issues raised at the group's December 1988 meeting.*

THE COMMON TENDENCY AMONG the general public to confuse archivists with anarchists may not appear so inaccurate when one considers the historic approach of archivists to description. We have often "followed our own lights" in determining what descriptive information should be provided about our holdings and how it should be presented. Wide variations occur among repositories, from the types of finding aids constructed (i.e., the components of each repository's "information system"), through the component parts or elements of information contained in each finding aid (i.e., the "data structure" of the finding aid), to the rules for entering information within each element (i.e., "data content" and "data value" standards).[1]

During its first meeting, the Working Group on Standards for Archival Description took note of the nearly vacant cells in its matrix for "internal data structures." While the group sensed that common practices have indeed evolved for preparing specific kinds of finding aids, none have been formally codified for formal adoption by the archival community. This paper analyzes prescriptions for practice as presented in several widely used manuals for archival description published during the past fifty years, drawing from them the implicit "data structures" recommended for finding aids most commonly used today in U.S. repositories. It also reviews recent developments, often instigated by automation, that have led to more explicit definitions of data structures.

The Importance of Data Structure Standards

Standards for data structures could help insure an adequate representation of information on archival holdings. A clearly defined set of data elements to be included in each type of finding aid would benefit the archival staff by providing a framework for producing consistent descriptive work. It would also assist new staff to adjust to the repository's descriptive practices. Since archival training is not required for many institutions, such guidance would be particularly useful in acclimating entry-level staff. For archivists taking another job in a new institution, it would lessen the need to learn "the way we do it here" and enable a quicker adjustment.

For researchers, common data structures could remove the need to spend hours learning how to access records each time they go to a new repository. Informal discussions with researchers about their experiences with archival finding aids indicates that they must factor this time into their research, and that they find the plethora of differing information to be frustrating and confusing.[2] Archives with a reputation for difficult access will continue to be underused.

Archivists themselves undertake activities for which predictable elements of information are required. Archivists in the Research Libraries Group (RLG) Seven States Project attempted to evaluate the usefulness of sharing scheduling and appraisal information on-line. One reason that an effective assessment could not be conducted, they found, was because the records created by project members lacked elements of information needed to make them comprehensible to others.[3] In addition to sharing information on appraisal and scheduling functions, archivists may find it

[1]For definitions and a discussion of these terms as used by the Working Group on Standards for Archival Description, see "Report of the Working Group on Standards for Archival Description," *American Archivist* 52 (Fall 1989): 454.

[2]These informal discussions have been held by description staff at the New York State Archives and Records Administration with a number of researchers who used the Archives' records over the past year.

[3]"RLG Seven States Case Study Assessment Report," Research Libraries Group, Palo Alto, CA, 1988. Unpublished project report.

useful to know whether records they are describing are related to holdings in other repositories. This issue is being pursued in particular by the National Archives and Records Administration's Intergovernmental Records Project, which is attempting to identify records held by various states and show how they relate to national records. Again, without common elements of information being recorded, it is difficult to determine the nature of records described by different repositories and whether they indeed have some common relationship.

Beyond the need to provide access to information about archival records more effectively, data structure standards are significant for facilitating automation. Frankly, the archives community is a small market for computer vendors or software developers. When archives have distinctly varied data structures, it is difficult and costly to develop software that meets each institution's individual needs. Few archives have the financial or technical capacity to develop their own systems, even if individually tailored systems were preferable. Because it is unlikely that vendors will develop such individually tailored systems for archives, development of archival software requires a common approach to data structures.

The best existing example of the value of data structure standards is the USMARC Format for Archival and Manuscripts Control (USMARC AMC). Since its adoption by the archival and library communities in 1983, it has made possible the inclusion of collection-level (and some item-level) descriptions of archival materials in national bibliographic utilities like RLIN and OCLC. But the lack of similar definitions for other types of archival descriptive records (finding aids) and associated data structures has meant that few vendors have developed systems for searching and providing access to anything much beyond the collection-level description available through USMARC AMC. Until common agreement is reached on more data structures, archival automation will be hampered.

Past Efforts to Establish Data Structures

Historically, archivists have made a number of efforts to identify the data structures they believed to be effective in various kinds of finding aids, especially guides, inventories, and collection-level descriptions. Rarely was there any concerted effort to push the recommendations forward for discussion or approval by the archival profession. The 1975 report of the SAA Committee on Finding Aids, published as *Inventories and Registers: A Handbook of Techniques and Examples* (1976), was a result of extensive analysis and deliberations but was never formally endorsed by SAA. The other manuals and guidelines listed in Figure 1 have often received wide acceptance and use, giving them the status of de facto standards in some cases, but they are not consensus standards in any sense. The data structures they describe are those that a particular author has distilled from his or her own view of "common practice" or are based on the practices of a specific repository with no attempt to address more general archival needs.

When one begins to search the archival literature, there is no dearth of monographs and articles suggesting data structures. The striking thing, however, is that each deals with the data structure for only one or two types of descriptive tools. In addition, the tools considered by each have their own overlapping but individual set of data structures. In no publication is a clear framework of related finding aids summarized along with data structures for each.

Current Developments

In addition to these earlier efforts that implicitly defined data structures for traditional finding aids, there are several current attempts underway to establish data

Figure 1

Manuals and Projects That Have Prescribed Data Structures for Archival Description

Manuals

Bordin, Ruth B., and Robert M. Warner. "Preparing Finding Aids." Chapter 4 in *The Modern Manuscript Library*, 50-68. New York: Scarecrow Press, Inc., 1966.

Bureau of Canadian Archivists. *Toward Descriptive Standards.* Ottawa: Bureau of Canadian Archivists, 1985.

Duckett, Kenneth W. *Modern Manuscripts: A Practical Manual for Their Management, Care, and Use.* Nashville: American Association for State and Local History, 1975.

Gracy, David B. *Archives and Manuscripts: Arrangement and Description*. Chicago: Society of American Archivists, 1977.

Hildesheimer, Françoise. *Guidelines for the Preparation of General Guides to National Archives: A RAMP Study*. Paris: UNESCO, 1983.

Hill, Edward E. *The Preparation of Inventories*. National Archives Staff Information Paper #14. Reprinted with slight revisions by author in *A Modern Archives Reader*, edited by Timothy Walch and Maygene Daniels, 211-35. Washington: National Archives and Records Administration, 1984.

Kane, Lucile. "Cataloging." In *A Guide to the Care and Administration of Manuscripts*, 2nd ed., 51-63. Nashville: American Association for State and Local History, 1966.

Lucas, Lydia. "Efficient Finding Aids," *American Archivist* 44 (Winter 1981): 21-26.

Muller, S., J.A. Feith, and R. Fruin. *Manual for the Arrangement and Description of Archives.* Translated by Arthur H. Leavitt, 1938. Reissue with new foreword by Ken Munden. New York: H.W. Wilson Company, 1968.

Network Development and MARC Standards Office. *USMARC Format for Bibliographic Data*. Washington: Library of Congress, 1988.

SAA Committee on Finding Aids. *Inventories and Registers: A Handbook of Techniques and Examples*. Chicago: Society of American Archivists, 1976.

Schellenberg, Theodore R. "Preparation of Archival Inventories and Guides." Chapter 14 in *The Management of Archives*, 219-39. Ithaca, NY: Cornell University Press, 1965.

Taylor, Hugh A. *The Arrangement and Description of Archival Materials*. With a contribution, "Les instruments de Recherches dans les Archives" by Étienne Taillemite. International Council on Archives, Handbooks Series Volume 2. New York: K.G. Saur, 1980.

Thibodeau, Sharon Gibbs. "Archival Arrangement and Description." Chapter 5 in *Managing Archives and Archival Institutions*, edited by James Gregory Bradsher, 67-77. Chicago: University of Chicago Press, 1989.

Current Projects

National Archives and Records Administration. "Archival Processing." Draft chapter 5 in *Office of the National Archives Procedures Manual, ARCHIVES 1400* (forthcoming).

Research Libraries Group, Government Records Project. "Minutes of Steering Committee Meetings." Research Libraries Group, Palo Alto, CA, 1988-89.

Research Libraries Group, Seven States Project. "Working Paper on Standards for RLIN Records for Records Schedules and Scheduled Records." Research Libraries Group, Palo Alto, 1987.

structures in very explicit terms, most of which are activities resulting from the development of automated description systems.[4] The original RLG Seven States Project, 1987-88, defined data structures for several descriptive tools that had not received much previous attention: agency records, scheduled records, and appraisal information records. A second expanded project, the RLG Government Records Project, is now attempting to define a range of types of archival descriptive records as well as the data structures for each. These include both authority records (e.g., agency history records, statutory records, and biographical records) and bibliographic records (e.g., accession-level records, full-level records, and record-system-level records).[5]

Several other developments also involve users of the RLIN system. The Archives, Manuscripts, and Special Collections Subcommittee on Descriptive Standards is pursuing definition of a full-level record for all AMC participants of RLIN.[6] The Historic Documents Inventory, a state-wide inventory of historical records repositories in New York State, defined a data structure for repository level records.

The National Archives and Records Administration has also begun to develop data structure definitions for information on holdings at the agency and series level, which it calls the "X" and "Y" Formats.[7]

While some of these archival descriptive tools and data structures have become conventions within a group or institution, they have not been submitted for professional review beyond the confines of the originators. Many are being followed by other institutions who have seen the tools and structures, and these may result in de facto guidelines or conventions for data structures over time. If these structures developed by individual repositories or small groups are adopted without being submitted to solid professional review, the danger of spreading idiosyncratic or inappropriate practice exists.

Elements Prescribed for Inclusion in Finding Aids and Other Data Structures

Figure 2 provides an overview of the data elements prescribed by the manuals and projects listed in Figure 1. It identifies those data elements that are frequently, occasionally, or rarely identified for ten types of finding aids and other data structures.

These sources have given the most attention to two descriptive record types: (1) the inventory, and (2) descriptions of collections or record series, both as part of inventories and as free-standing entities (called collection/series-level description records). They direct considerably less attention to several other types: guides, calendars, card catalogs, special indexes, accession-level records, and, more recently, agency history records, general schedule records, and scheduled item records.

The inventory. Several archivists give attention to what is variously called an inventory, a register, or simply a finding aid. Although there are some differences between the three, the structure for all are defined similarly. Edward Hill provides the most detailed data structure in his National Archives Information Paper #14. This may be because Hill was prescribing a structure

[4]For information about projects discussed in this section, see the project reports listed in Figure 1, in addition to the reports cited in the following footnotes.

[5]RLG Government Records Project, "Online Record Types for Government Records," Research Libraries Group, Palo Alto, July 1990. Unpublished draft.

[6]"Bibliographic Standards for the Use of the AMC Format," Research Libraries Group, Palo Alto, July 1990. Unpublished draft.

[7]"Archival Processing," draft chapter 5 in *Office of the National Archives Procedures Manual, ARCHIVES 1400* (Washington: NARA, forthcoming).

Figure 2

Information Elements Prescribed for Specific Types of Finding Aids and Other Data Structures

	ELEMENTS OF INFORMATION PRESCRIBED		INVENTORY	SERIES/COLLECTION WITHIN INVENTORY	COLLECTION/SERIES-LEVEL DESCRIPTION	GUIDE	CARD CATALOG	ACCESSION-LEVEL	AGENCY HISTORY	GENERAL SCHEDULE	SCHEDULE ITEM	APPRAISAL INFO
	MOST COMMON FORM USED IN WRITINGS ON DESCRIPTION	EQUIVALENT TERM IN SAA DATA ELEMENT DICTIONARY (1983)										
INFORMATION ABOUT THE CREATOR OF THE RECORDS	Name of records creator	CREATOR; OFFICE OF ORIGIN	●	◐		●					●	●
	Main entry			◐	●		●	●	●			
	Organizational history/ biography	BIOGRAPHY/HISTORY	●		◐	●			●			
	Predecessor and successor agencies		◐						N			
	Major functions/activities	FUNCTION	●		◐				N			
	Legal authorizations	AUTHORIZATION	◐									
	Major subdivisions	INSTITUTIONAL DIVISION	◐						N			
	Creator's recordskeeping practices		◐									
	Names of agency heads	AFFILIATION; AFFILIATION ROLE	◐									
	Dates related to main entry	VITAL DATES					●					
PHYSICAL CHARACTERISTICS OF THE RECORD	Quantity/volume of records	MEASURE; QUANTITY; PACKAGING UNIT; SIZE OF PACKAGING UNIT	◐	●	●	●	●	●		●	●	●
	Arrangement statement	ARRANGEMENT	●	●	●	◐						
	Type of material; form/ genre terms	FORM; GENRE; MEDIUM	◐	●	●					●	●	●
	General material designator				○							
	Location					◐	●			●		

● = Frequently Identified Elements
◐ = Occasionally Identified Elements
○ = Rarely Identified Elements
N = Required by the National Archives and Records Administration (Agency History Records Only)

Figure 2 *(Continued)*

	ELEMENTS OF INFORMATION PRESCRIBED		INVENTORY	SERIES/COLLECTION WITHIN INVENTORY	COLLECTION/SERIES-LEVEL DESCRIPTION	GUIDE	CARD CATALOG	ACCESSION LEVEL	AGENCY HISTORY	GENERAL SCHEDULE	SCHEDULE ITEM	APPRAISAL INFO
	MOST COMMON FORM USED IN WRITINGS ON DESCRIPTION	EQUIVALENT TERM IN SAA DATA ELEMENT DICTIONARY (1983)										
INFORMATION ABOUT THE INTELLECTUAL CONTENT OF THE RECORD	Title (of record group, series, item)	TITLE OF RECORD UNIT	◐	●	●	●	●	●	●	●	●	●
	Previous titles				○							
	Inclusive dates	INCLUSIVE DATES	●	●	●	●	●	●	●		●	●
	Bulk dates	BULK DATES	◐		○							
	Scope & content note	ABSTRACT; SCOPE AND CONTENTS NOTE		●	●	●	●	●		●	●	●
	Provenance	PROVENANCE		◐	◐							
	Research strengths		◐									
	Significant items (having intrinsic value)			◐	◐							
	Genealogical/ organizational charts		◐									
INDEXING TERMS/ACCESS POINTS	Identifying number	CONTROL RECORD IDENTIFIER; LOCAL CATALOG/CONTROL NO.		◐	○							
	Principal subjects/topics	SUBJECT CATEGORY; TOPICAL SUBJECT REFERENCE	●	●	◐					●	●	●
	Personal name references/significant correspondents	PERSONAL NAME REFERENCE	◐	◐	○			N				
	Geographical references/Place names	GEOGRAPHIC REFERENCE		◐	○			N				
	Corporate name references	CORPORATE/ CONFERENCE NAME REFERENCE			○							
	Index											

for one specific institution—the National Archives—with a common type of record, i.e., government records. T. R. Schellenberg, Ruth Bordin and Robert Warner, the SAA Committee on Finding Aids, and David Gracy identify many similar elements, but in much less detail than Hill.[8] Separate sets

[8]See Figure 1 for citations to the works referred to in this and following paragraphs.

Figure 2 *(Continued)*

	ELEMENTS OF INFORMATION PRESCRIBED		INVENTORY	SERIES/COLLECTION WITHIN INVENTORY	COLLECTION/SERIES LEVEL DESCRIPTION	GUIDE	CARD CATALOG	ACCESSION LEVEL	AGENCY HISTORY	GENERAL SCHEDULE	SCHEDULE ITEM	INFO APPRAISAL
	MOST COMMON FORM USED IN WRITINGS ON DESCRIPTION	EQUIVALENT TERM IN SAA DATA ELEMENT DICTIONARY (1983)										
INFORMATION AFFECTING USE OF THE RECORD	References to related records	ASSOCIATED MATERIALS; RELATED MATERIALS	●						●	●		
	Box/folder lists	ITEM NUMBER; PACKAGING UNIT	◐		○			●				
	Bibliographic citations	BIBLIOGRAPHIC REFERENCES	◐		○							
	Microfilm publications available	COPY IDENTIFICATION		◐	◐							
	Finding aids available	FINDING AIDS		◐	◐							
	Published descriptions of the collection	BIBLIOGRAPHIC REFERENCES			○							
	Reproductions available	COPY IDENTIFICATION		◐								
	Technical Access requirements	TECHNICAL ACCESS REQUIREMENTS			○							
	Restrictions	TERMS GOVERNING ACCESS; TERMS GOVERNING USE			◐							
	Copyright information				○							
	Retention & disposition note	RETENTION REQUIREMENTS							●	●	●	
INFORMATION ABOUT THE FINDING AID	Cutoff date of finding aid	DATA ENTRY TIME; TIME OF ACTION	◐									
	Name of preparer/author of finding aid	ACTION AGENT			○				N			
	Name of reviewer of finding aid	ACTION AGENT			○				N			
	Acknowledgements		◐									

of data elements are identified for the overall inventory itself and for descriptions of individual series or collections within the inventory.

Collection/series-level description record. The descriptive record type for which the largest number of articles or monographs prescribe data structures is the collection/series-level record. This is not surprising because it is the descriptive record most necessary to gain control over and provide access to archival records. Structures are provided in earlier literature, including work by T. R. Schellenberg; the Dutch archivists Muller, Feith, and Fruin; Lucile Kane; and Edward Hill. Structures have been prescribed in more recent years by Lydia Lucas, the SAA Committee on

Finding Aids, the Bureau of Canadian Archivists, and the National Archives.

Guides. Although it is a common type of descriptive record, considerably less attention is given to data structures for guides. Françoise Hildesheimer devotes the most attention to it. Schellenberg also offers a bit of advice on data structures for guides. More recently, the SPINDEX automated system has provided a de facto structure for guides by the nature of the data elements it defined. Most commonly, archivists seems to look at actual guides created by other repositories and adapt structures given in them.

Card catalog. Reliance on the card catalog as a finding aid has diminished, but guidelines have been provided in the past by Lucile Kane and by Ruth Bordin and Robert Warner, both in publications directed at manuscript repositories. They agree almost exactly on the elements to be included.

Accession-level record. Lydia Lucas has recommended a structure for an initial record created at the time of accessioning to be filled out in more detail at a later time.

Agency history record. Recent developments in automation, and efforts to accommodate government record information on the RLIN system have led members to define an agency history record. This general record, one for each agency, can then be linked to each of the records for individual series created by that agency. This circumvents the need to enter the same historical note repeatedly in potentially hundreds of series-level descriptions. The structure has been adapted by the National Archives, and interest in it has been expressed by various other institutions. This record type does not come, however, out of common archival practice. Instead, it was necessitated by the inability of the RLIN system to provide searchable authority files. It points to the need for defined authority type records for archives.

Schedule and appraisal records. Three other types of descriptive records were defined by government archivists in the RLG Seven States Project. Government records repositories often produce general schedules, item schedules, and appraisal information, but have seldom discussed them in the context of archival description. Until recently, archivists have had little control over records management activities or else have not taken a proactive role in attempting to control the type of information produced during records management activities. The Seven States Project has defined data structures for both general and specific-item schedules, as well as for records containing appraisal information.

The Need for a General Archival Information System

Until recently, discussions of finding aids and related data structures have been confined to a conservative range, predominantly inventories and collection/series-level description records. One might conjecture endlessly (and with some perverse delight) over the reasons for this failure to investigate the actual types that have been—or might need to be—created. In the last five years, significant forays have begun to look at and define a greater range of archival descriptive record types, those being agency history, schedule, and appraisal records. These efforts by government records archivists need review by the profession at large before conclusive statements can be made about them first, as effective types of finding aids, and, then, about the detail of their data structures.

Development of viable standards for data structures cannot proceed in the absence of standards for archival information systems. An archival information system is a rational framework relating a group of descriptive record types that contain the information needed to carry out archival

functions. It may include both manual and automated record types and will most likely involve some variations among repositories based on local needs.

To date, archives have not clearly defined the complete range and relationships of descriptive record types for traditional, manual systems. While certain pieces of a potential system can be found in the published discussions of record types analyzed above, there has been no systematic effort to define each type, its purpose, its relationship to other descriptive record types, and how it relates to specific archival functions such as appraisal, accessioning, description, and reference. These record types are the framework for recording and locating information on archival holdings, yet they have not received solid definition and analysis. We have relied on the "organizing" skills of individuals within a repository to determine ad hoc what the framework for an archival information system will be. The result is a proliferation of somewhat idiosyncratic descriptive records types which both archivists and users must adjust to each time they change repositories. Existing systems may meet the perceived needs within a given repository, but other potential means of access or types of information may go unexplored because of the confinements of that local system or the staff's lack of comprehensive understanding of archival functions and the relationships of the information created by these functions.

The introduction of automation into archives further complicates the general problem caused by the lack of definition of the descriptive record types composing an archival information system. While automation offers new opportunities for eliminating redundancies in recording information and providing more facile retrieval, it complicates the definitions of descriptive record types. The computer does not simply duplicate manual record types in an on-line mode. Archivists need to turn their attention to the relationship of archival functions to manual and automated descriptive record types in order to create effective, efficient archival information systems.

A carefully structured general archival information system standard can and should be defined. Local circumstances may require adaptations to accommodate certain types of materials or repositories, varying functions carried out by each repository, or particular automated systems in use, but they should be able to fit within a common conceptual framework. The archival profession badly needs such a framework if further work on description standards is to proceed. Data structures cannot be defined effectively without first defining the components of an archival information system. It makes little sense to provide standards that define the elements to be included in a guide, a register, or an on-line catalog if the interrelationships among those record types are not standardized.

First of all, the archival profession should identify what other archival descriptive record types are being, or need to be, created. Second, it must address the need for data structures relating to special formats or forms of material such as maps, historical photographs, machine-readable records, architectural drawings, sound records, or archival films within a traditional collection/series-level description structure. Microforms or other multiple versions of records also require special attention regarding descriptive data structures.

Existing archival descriptive record types and data structures have focused almost exclusively on bibliographic information about the content and context of records. An archival information system could be enriched by the inclusion of information on processing actions such as weeding or rearranging, or conservation actions such as deacidifying, mending, flattening, and fumigating. Management activities also provide significant data elements such as donor,

method of acquisition, and date of acquisition. Serious consideration needs to be given to whether these types of information should be considered part of the archival descriptive record and, if so, what the accompanying data structures should be.

Issues for the Working Group to Consider

ISSUE #1: Defining record types and data structures. Descriptive record types for an archival information system need to be defined, followed by definition of data structures for each record type. There is clearly a need to bring together the plethora of recommendations and thoughts on finding aid types or tools. Fifty years of archival literature becomes almost depressing in its repeated discussions of the tools and data structures archivists should be constructing. Bits and pieces are reiterated again and again, but no overall framework for the types of records in an archival information system is provided.

ISSUE #2: Broadening current initiatives. As described above, a variety of groups and institutions are working to establish descriptive record types and data structures based on the interests, needs, and motivations of a limited number of institutions. For maximum effectiveness and applicability, the work done by these groups should be reviewed and extended to meet the needs of the profession at large in order to address the types of descriptive tools, the needs of special materials, and the concerns of institutional types not represented by their groups.

ISSUE #3: Identifying leadership. Who should take the lead role in defining areas of needed development? Potential actors include the Working Group, some successor group, and the Society of American Archivists through its Committee on Archival Information Exchange or its Description Section.

ISSUE #4: Identifying participants. Who should undertake the development of data structure guidelines, conventions, and technical standards? Potential actors include the Working Group, some successor group, SAA through its Committee on Archival Information Exchange or its Description Section, groups of institutions with common concerns such as the RLG Government Records Project members, and individuals commissioned to survey the profession and prepare discussion summaries.

ISSUE #5: Obtaining resources. How can resources be obtained to accomplish this work? Potential sources include federal funding agencies, private funding agencies, computer vendor systems, institutional funding, and SAA or other professional organizations.