Archives and Museum Informatics
Technical Report

ISSN 1042-1459 No. 6

Archives and Authority Control

Proceedings of a seminar sponsored
by the Smithsonian Institution
October 27, 1987

edited by Avra Michelson

Originally published as Archival Informatics Technical Report
Vol. 2 No. 2, Summer 1988
Reprinted 1989

Table of Contents

Acknowledgements

Many individuals that I would like to acknowledge contributed to this publication. The SIBIS–Archives Users Group Descriptive Standards Committee, comprised of Diane Vogt O'Connor, Judy Throm, and myself, and chaired by Harry Heiss, organized the seminar and thus made this volume possible. Richard Szary contributed to the conceptual development of the seminar. More than 100 Smithsonian staff, including archivists, registrars, and related information professionals set aside pressing work to participate in the seminar. The Smithsonian Institution's Office of Information Resource Management generously underwrote seminar expenses. Harry Heiss compiled the extensive bibliography that appears at the end. I extend a special thank you to Susanne Peckham for her superb editing. Wanda Perkins heroically handled administrative tasks related to publication. Ralph Walker made "camera–ready copy" a reality. David Bearman provided empathic guidance throughout the editorial process.

Introduction

Descriptive Standards and the Archival Profession

Avra Michelson

A mere five years ago, archivists began using the MARC AMC format to share information in national networks.[1] Concurrent with this national endeavor, archivists at the Smithsonian Institution began developing an online local system, the Smithsonian Institution Bibliographic Information System (SIBIS). A MARC catalog that operates on Geac hardware and software, SIBIS encompasses three separate databases: one devoted entirely to archival and manuscript materials, one for library materials, and one for research files. Even before the system became operational in 1984, Smithsonian archivists met to participate in defining system requirements and shaping development. Having established an online database, in 1986 the SIBIS-Archives users group convened a committee to define and implement descriptive standards that would improve access to information held in the archives catalog.[2]

The committee was charged with recommending standards for description that would serve the seven different archival repositories comprising the database.[3] SIBIS participants included various types of repositories (e.g., institutional archives as well as special collections), with a range of subject concentrations (e.g., anthropology, art, and folklife), and specialized media (e.g., film, sound recordings, photographs, text, and so forth). The committee quickly focused its efforts on authority control as a mechanism to regulate content that appears in headings (e.g., expressions that represent a name or concept that serves as an access point to a record). While recognizing the importance of establishing guidelines for choosing forms of headings and prescribing input styles, to standardize these areas would prove of little value if the headings themselves were not controlled. Although authority control served as the point of departure, the committee recognized that a comprehensive set of cataloging standards would have to be adopted to achieve their goals.

The SIBIS-Archives Descriptive Standards Committee organized the *Archives and Authority Control* seminar, held October 27, 1987, to prepare both the committee and the users group to define an authority control system that would improve retrieval in the bibliographic catalog. Like many in our profession, we felt ill equipped to consider this task without substantial education, especially since a number of authority files would be necessary to improve access to the varied materials cataloged in the archives database. The conference both introduced participants to conventional library uses of authority control and considered the development of independent authority reference files as the logical outgrowth of archival description. Probably most important, the conference articulated why the development of archival descriptive standards deserves prominence on the professional agenda.

Standards are agreements that guide the practice of professionals within a discipline.[4] Archival administration is guided by standards in a number of areas. For instance, archivists generally accept that materials require storage within a specific range of temperature and humidity, off-the-floor shelving, and preservation in acid-free containers. These constitute very general preservation standards. In a slightly different arena, descriptive standards provide rules that promote consistent intellectual control and access to bibliographic materials. Descriptive standards, while useful in establishing consistent headings for manual systems, play the cornerstone role in successful retrieval in online bibliographic networks. Archival descriptive standards are best understood on a continuum that includes three categories: data structures, data contents, and data values.[5]

Data structures are "buckets" or "containers" that

hold data. Since every data base design includes a data structure, and since the information containers used from system to system can differ, many different data structures exist at any one time within a profession. But for the purposes of information exchange, a profession must adopt a common data structure. Currently in our profession, archivists use the MARC AMC format as the uniform format for information exchange. As a data structure, the MARC AMC format organizes the categories of information contained in a record.

The second category of standards refers to data content standards or the guidelines that provide rules for formulating words, terms, or phrases entered in data structures. The *Anglo-American Cataloging Rules*, 2nd edition (AACR2) is an example of this type of standard. A portion of AACR2 establishes rules for choosing and constructing names: Should you choose a pseudonym or a given name as the authorized heading? Do you choose a correspondent's maiden name or married name? Steve Hensen's manual *Archives, Personal Papers, and Manuscripts* (APPM) represents a related type of data content standard, defining a basic set of cataloging rules for describing archival and manuscript materials.

Finally, there is the third category of descriptive standards, data value standards. They are the subject of this proceedings. Often referred to as authority files or controlled vocabularies, they may appear as simple lists or as thesauri with syndetic structure. Authority files contain information that allow catalogers to transform a concept, name, or topic into an actual value (word, term, or phrase) to be entered as an "authorized" or "established" heading into a particular MARC field. Authority control standards provide guidance in the choice of data values by authorizing certain terms. For instance, using the *Library of Congress Subject Headings* (LCSH) as a controlled vocabulary, the correct terminology for the concept of the Vietnam war is "Vietnamese Conflict, 1961-1975". Similarly, the LC subject heading for medical doctors is "Physicians." The *Art and Architecture Thesaurus* provides another example of a value standard as do specialized thesauri or controlled vocabulary lists that members of specific disciplines agree to follow.

Library practice identifies catalog patrons, rather than librarians, as the intended beneficiaries of authority controls. As we are only now beginning to study the retrieval needs of our users, archivists have not yet defined clear criteria by which to measure a successful catalog. But since our profession has chosen to share information through participation in bibliographic networks, it is appropriate to consider what librarians have identified as the purpose of their catalogs and to evaluate to what extent this has relevance for archival practice.

Jackie Dooley relates that the purpose of a library catalog is to enable a searcher to find any known item, by author or title, as well as groups of items, such as all the works of an author, all the editions of a work, or all works on a particular subject. The ability to retrieve a particular item as well as a group of like items requires consistency in the choice and form of all headings. Dooley cites three key elements that promote consistency to meet these objectives: the ability to distinguish names, show relationships, and document decisions. The primary way library catalogs attempt to achieve consistency is through the use of authority controls. Authority control refers to the regulation of terminology or concepts used as headings in catalog records. This includes the process of determining the form of a name or of a concept, determining all the necessary cross-references that will be referred to that heading and from other established headings, and documenting decisions made during authority work.

Authority control aids consistency through the use of controlled vocabularies, which collocate like materials and link related but separate materials. By itself, however, authority control cannot improve retrieval in online catalogs. But used in conjunction with data content standards and a standard format for data exchange, authority control provides the best-known basis by which users can evaluate their searches. As Dooley points out, in online catalogs comprised of records constructed with a disregard for standards, users have no way of knowing if a search resulting in "no hits" is truly a search failure, or if they simply should continue searching synonymous terms or other forms of names. In the absence of standards, how do researchers assess their search results? The adoption of cataloging rules and authority controls standardizes the way archivists create headings, which in turn enables researchers to reasonably assess the results of their searches in online catalogs.

Although the primary reason to adopt authority control and descriptive standards is to maximize the ability of researchers to retrieve primary source materials pertinent to their work, the adoption of standards promises to benefit archivists in the management of descriptive programs as well. Applying a standard set of rules to guide the form and choice of headings and using controlled vocabularies and authority files to guide the choice of values ultimately *simplifies* rather than *complicates* the descriptive process.

Adoption of standards means that equivocal headings are considered only once with a conclusive decision reached for that name or topic. It means that each repository is freed from the intensive labor needed to develop a comprehensive set of bibliographic, standards. Rather, existing standards

such as *Library of Congress Name Authorities* and *Library of Congress Subject Headings* can be used to control personal names, corporate names, and the majority of topical headings, which allows archivists to focus on the development of specialized descriptive tools, such as controlled terminology for form and genre, physical characteristics, occupations, and function terms. Finally, adopting standard authority controls alleviates a significant portion of the labor involved in controlling headings; where available, archivists can refer to authority files maintained by other bodies and develop authority records only for those headings that do not already appear in existing files. Although implementation of standards undoubtedly translates into an increased work load for archivists in the short term, it is nonetheless a necessary outgrowth of online information exchange that will, in the long term, maximize the use of limited archival resources.

The library profession uses standard conventions, including the *Library of Congress Name Authorities*, the *Library of Congress Subject Headings*, and the *Anglo-American Cataloging Rules*, 2nd edition, to guide cataloging of materials. These conventions allow librarians to describe thousands of copies of the same item at reduced costs through copy cataloging — sharing descriptions of nonunique items among repositories. Instead of each library creating a catalog record for its particular copy of *Gone With the Wind*, libraries copy and then customize an existing description within a database. Shared cataloging was initiated first as a manual process through the Library of Congress card distribution project. It continues in the automated environment with the copy cataloging of MARC records. The library community achieves a compelling financial benefit through standardization of descriptive data elements, rules that control forms of headings, and values that serve as access points to records.

Archival description has matured in a significantly different professional environment. Aside from microforms, archivists describe primarily unique materials that require original cataloging. Until the development of the MARC AMC format, and the resulting capability to exchange information, the profession has had little impetus to develop and use descriptive standards. The notable exception is Steve Hensen's *Archives, Personal Papers, and Manuscripts*, adopted by the Research Libraries Information Network (RLIN) Archives and Manuscripts Control (AMC) participants, which most closely approximates an archival standard. Although archivists have developed few standards of their own, use of descriptive standards with primary source materials is no longer debatable — rather, it is required if we are to retrieve material with reasonable success from the databases under formation.

Information exchange in an online environment presents a challenge to the archival profession's traditional descriptive programs and processes used to provide access to materials of enduring historic value. The implementation of network and local authority control warrants a central role in descriptive programs. Before, however, these systems can be adopted, a series of questions must be posed concerning current descriptive policies and procedures.

As a profession, we still have not clarified nor agreed on the purpose of online catalog descriptions: Are we creating catalog records for other archivists who will continue to act as intermediaries at the online catalog, or are we creating databases for direct access by users? That these questions arise emphasizes how very little we know about our users — who they are; how they approach research with primary source materials; what constitutes significant points of access; and what kind of structured vocabularies and authorities are needed to support a variety of access approaches. The response to these questions, although essential to the development of client-centered databases, has received insufficient consideration to date. New policies and procedures that address the relationship between use, description, and access are called for. As a logical outcome of this process, the distinction of description and reference into functionally separate activities may be found to no longer serve the archival mission.

The process of examining description policies should result in improved descriptive practice. Although transformation of our descriptive practice is underway, it has not been an entirely easy or successful process. My study of RLIN AMC participants found that the lack of consistency among archivists in the assignment of topical index terms constitutes a major stumbling block to retrieval of information, and that indexing inconsistency, the lack of adequate authority controls, and problems with the use of LCSH all inhibit our ability to retrieve materials in bibliographic utilities.[6] Training in the use of library conventions with bibliographic networks should temper these problems as a more standard practice develops.

Finally, assuming that archivists have determined the purpose of online catalogs and are trained in the use of library standards, these standards must then be evaluated and modified: To what extent are they adequate for use with primary source materials and to what extent do they pose problems for modern archival description. To adequately evaluate library standards however, archivists as a profession must be both well trained in the use of those standards and in information science techniques pertinent to archival information exchange. Evaluation of the

usefulness of library standards entails assessing the extent to which established guidelines and controls apply to use with primary source materials, recommending changes to oversight bodies responsible for maintaining specific standards, coordinating the development of new controls where necessary, and overseeing the development of archival guidelines for use with library-generated standards.

By itself, authority control does not adequately improve retrieval in local and national networks. But in conjunction with other bibliographic standards, the implementation of authority controls achieves significant results. Archivists in their descriptive practice, therefore, should follow *APPM*, adopt *AACR2* rules, use the *LCNA* and related authority files, and adopt the use of controlled vocabularies, such as *LCSH*, to control topical access to materials when creating bibliographic catalog records.

The papers that comprise this proceedings address many of the key issues that surround implementation of authority controls with primary source materials. Jackie Dooley begins with a conceptual framework for understanding authority control and its use with archival and manuscript materials. Tom Garnett then shares his experience in developing authority controls for the library database within the SIBIS catalog, emphasizing the real-world trade-offs involved in implementation. Next, Marion Matters outlines the process by which a prototypic authority system was integrated into the ongoing descriptive program at the Minnesota Historical Society. Lisa Weber follows with a challenge to the archival community to expand conventional library usage of authority control to include the development of independent biographical and organizational reference files, and she assesses the place of this work on the current professional agenda. In the final paper of this collection, Richard Szary presents the technical requirements for the archives database on SIBIS, identifying the function of each requirement and the current status of software development. The concluding portion of the proceedings includes an hour-long transcript of a discussion led by David Bearman, in which the speakers respond to audience questions with concrete advice on how to begin implementation of an authority control system.

The use of authority control and the adoption of descriptive standards to facilitate information exchange will necessarily transform current archival descriptive practices. Since the primary mission of archivists is to make materials accessible for use, in the age of automation anything short of grappling with the issues that allow us to share information in networks shirks that mission. The proceedings demonstrated that the development and use of descriptive standards, which include authority controls, presents some of the most important work facing archivists as it directly challenges our professional commitment to provide access to materials and the extent to which we are willing to make primary source materials housed in our repositories available to a larger audience. The development and use of standards is not an end in itself, but rather the vehicle that balances archival methods of description with the retrieval needs of our users — and certainly a greater balance is in order.

Many ideas in this paper were presented on April 29, 1988, at the "Library Descriptive Standards" workshop in Chicago sponsored by the Society of American Archivists. Lisa Weber and David Bearman commented on several drafts of this paper and made significant contributions to its development. Richard Smiraglia and Richard Cox provided useful comments on an earlier version. The author is grateful for this assistance.

NOTES

1. Archival programs at Cornell University, Stanford University, and Yale University comprised the participants in the 1983 RLIN AMC pilot project.

2. The SIBIS-Archives Descriptive Standards Committee was chaired by Harry Heiss of the National Air and Space Museum Archives, and included Diane Vogt O'Connor of the Photo Survey Project of the Smithsonian Institution Archives, Judy Throm of the Archives of American Art, and myself, representing SIBIS administration.

3. SIBIS-Archives database participants include the Archives of American Art, Human Studies Film Archives, National Air and Space Museum Archives, National Anthropological Archives, National Museum of American History Archives Center, Office of Folklife Programs Archives, and the Smithsonian Institution Archives.

4. David Bearman supplied this definition of standards.

5. I am grateful to David Bearman and Lisa Weber for suggesting this framework to me for understanding descriptive standards.

6. Michelson, Avra. "Description and reference in the age of automation." American Archivist, Vol. 50 (Spring 1987), p. 192-208.