Development of Authority Control Systems Within the Archival Profession

Lisa B. Weber

"Authority Control." Repeat that phrase to the person on the street and see what kinds of looks you get. When I was working at the State Historical Society of Wisconsin in Madison I had Mary Ghikas' book[1] on the shelf above my desk. For those of you who've never seen it, the book has a severe black cover with bold white letters on the spine spelling just the title "Authority Control" and not including the subtitle "The Key to Tomorrow's Catalog." A friend of mine once asked me about the book, in such a way that I could tell he was worried that I was either involved in some sort of paramilitary organization or sadomasochistic sexual practice. I explained it was neither, but that's what the phrase "authority control" implies to the uninitiated.

After the papers presented thus far, I feel safe in saying, "Welcome to the initiated." It may not surprise you to know that until recently most archivists (especially those with no library science background) would fall into the uninitiated category. But that trend is changing, fairly rapidly I might add, because of the strong incentive for archivists to use controlled vocabularies and descriptive standards. Although the need has always existed inhouse - even if many archivists did not recognize it - the force behind this change has been the use of automated applications and in particular the desire to exchange information via the MARC format.

It is my task to talk about the development of authority control systems in the archival profession. This encompasses defining archival requirements and adapting conventional authority control systems to archival needs, summarizing archival literature on authority control, and placing authority control in the context of the profession's agenda. As I do not want to dwell on how archivists can just adapt library conventions, I also will focus on how archivists can go beyond adaptation. This includes expanding authority control concepts in ways that may not only improve access to historical materials but may succeed in bringing archivists (perhaps reluctantly) into the mainstream of the information community.

Before we can examine how to adapt conventional authority control to archival needs, the first question we must ask is:. "What are archival needs?" Clearly, in the library world, authority control is directly tied to the purpose and function of the library catalog. As first codified by Charles Cutter in 1876 and restated in the Paris Principles of 1961, the library catalog serves to (1) locate a particular known item and (2) gather together or collocate a category of works. Librarians have assumed that a large percentage of searches are for known items (i.e., where a user approaches the catalog with an author or title in mind). But catalogers also assign subject headings to bring together items with the same intellectual characteristics. The success of the library catalog depends upon using a standard set of headings embodied in the concept of authority control, in addition to a standard set of transcription rules. All these standards are strongly supported on a profession-wide basis.

This kind of clearly articulated, purposeful statement of cataloging requirements and authority control relationships does not exist in the archival community. We can best answer the previously posed question ("What are archival needs?") with "We're not sure." The role of cataloging in archives is unclear, although the use of the MARC AMC format is forcing archivists to think more about this relationship.

Although to my knowledge this belief has never been tested, archivists assume that most people approach archives without having a specific item in mind. Since archival material is not published, the traditional bibliographic access points of author, title, and even subject are only somewhat useful. Archival materials do have creators that are analogous to authors, but most of the titles of archival materials are supplied by archivists. And some of the problems with subject access stem form the fact that archival materials are the unconscious product of activity

rather than intellectually cohesive efforts from authors. Archival materials are not "about" something or someone as much as they are "by" some corporate body or person.

The Paris Principles characterize the catalog more as a tool of bibliographic description than of reference access, placing greater emphasis on the physical rather than the intellectual attributes of items. In a paper delivered at the 1981 SAA annual meeting in Berkeley, California, Steve Hensen states that:

> Unfortunately this move was in exactly the wrong direction for manuscripts and their description. The Paris Principles represent as clear a statement of what the principles of manuscripts cataloging should NOT be as can be found. For the cataloging of manuscripts is primarily a question of intellectual access with the catalog serving as a tool of reference and research and not of bibliographic description....The users of manuscript catalogs are less interested in the physical characteristics of the manuscript than in their context and content.[2]

As Hensen so aptly states, the focus on bibliographic description in library catalogs (as expressed in the Paris Principles) is limiting to archivists. Archivists want to create catalogs that go far beyond describing the physical characteristics of the records to ones that provide users with as much information as possible concerning the provenance (i.e. the *context)* of the materials as well as information about the intellectual *content* of the records. Archivists and manuscripts curators want to create catalogs that are tools of reference and research. They can do this by using and expanding on the concept of authority control. Authority control can provide the context and the content, and thus overcome these descriptive limitations.

Archivists need to take several steps to do this: first, additional authority-controlled access points are needed for archival material; second, the existing conventional library practice of establishing headings, documenting decisions, and creating cross references must be expanded to include the building of reference files as important source files in themselves; and third, these reference files must be available to researchers.

Since use of the MARC AMC format enables primary source descriptions to be integrated with descriptions of other materials located in libraries, it is important for archivists to continue to use traditional access points. But, at the same time, archivists must develop as well authority-controlled vocabularies for other, more archivally oriented access points.

These more archivally oriented access points include function, genre, and form of material. Access by function provides a means for identifying information based on the intent or purpose with which archival (i.e., corporate) records were created. Identifying the function of a record series answers the question of "why" the records were created, since corporate bodies come into existence for identifiable purposes. For example, one function of a corporate body may be to inspect (as in the case of state government agencies) or to lobby (as in the case of professional associations). One advantage to assigning "function" access points is the possibility of collocating similar materials by function of the creating body. The names of the agencies or their position in an organization structure take on secondary importance as the critical access point would be function.

Other access points applicable to archival materials include "genre" and "form of material." These categories enable archivists to answer the question, "How was the information recorded?" and to retrieve by such terms as daybooks, syllabi, or hymns. In the MARC AMC format, there are fields for all the categories just mentioned.

As interesting as increasing the number of retrievable access points for archival materials is the notion of access via the provenance approach. Richard Lytle, David Bearman and Richard Szary argue that the provenance approach — using information about the history, structure, and activities of persons and organizations that created the materials — is a superior method of gaining access to archival materials compared with the assignment of subject headings.[3] How often have reference archivists transformed a topical question into an organizational activity statement by asking what organization or what part of an organization could have been engaged in activities that would produce the kinds of information a patron seeks. Or how often have archivists tried to compile the personal papers of members of specific organizations, ethnic backgrounds, or occupational areas? If authority records existed for organizations and people that included this kind of information that we have traditionally put in our historical/biographical notes, we would, in effect, create our own reference files that could be searched separately but linked to descriptions of the records or the records themselves.

Files like these need not be limited to people or corporate bodies but could be created for geographic entities, culture or linguistic groups, or occupational areas. In fact, Bearman argues that much "authority data" already exists, having been created by the public sector in the form of the federal *Government*

Organizational Manual, the *Federal Register*, and the various state and local government manuals. In the private sector, there are the various biographical dictionaries like *Marquis's Who's Who*. It would be a substantial task to build bridges between this kind of authority data and archival authority control systems.

Implementing the ideas of Bearman and Szary requires the development of non-hierarchical architectures that do not attribute a "privileged position" to any file. This means that one file should not be more important than any other file. For example, librarians create authority records to help provide access to bibliographic records. In essence, the bibliographic record holds a more important or privileged position. Bearman and Szary argue that the authority record does not exist only to serve the bibliographic record but is a peer or equal to the bibliographic record. It is a reference in its own right. We need to create, to use Bearman's phrase, multiple peer data bases, for "one man's authority file is another man's data base proper."[4]

Though all of this is exhilarating, it is time to take our heads out of the clouds and talk about realities. The fact is that developing and maintaining authority control systems, even simple ones, take resources. I do not mean just money, although money is certainly needed. Of equal importance is expertise. The kinds of authority files I've discussed are certainly far more sophisticated and complex that the authority files that librarians have been creating for quite some time. Cost analysis figures quoted by R. Bruce Miller, although now almost 10 years old, are quite sobering.[5] But Miller also makes the compelling observation that "librarians have not truly recognized the expense of not creating rigorous authority control." The archival community, unfortunately, does not have comparable cost analysis figures. Therefore, it is very difficult to evaluate the price we pay.

What about the real world? Many of you know that there is tremendous interest in the MARC AMC format and that more than a handful of archival repositories are using it. Archivists are currently following standards to describe materials and using controlled vocabularies when constructing access points. Marion Matters has just discussed how the Minnesota Historical Society is coping with authority control. The Minnesota Historical Society is also a part of the Research Libraries Group's (RLG) Seven State Archives Project. Under this project, a group of seven states are currently putting descriptions of state and local government records into RLG's Research Libraries Information Network (RLIN). That group of archivists is also creating a functions vocabulary to provide access points to the records descriptions. In addition, those states are creating two kinds of records, one type called agency history records, which are really administrative histories or authority data about the creating bodies. The agency history records are then linked to descriptions of the actual archival material as described in series records. The implementation is not as "clean" as it could be since one data base holds both the bibliographic and authority records, but RLG was willing to make this compromise for the sake of the project.

Another important archival authority control activity is that Harriet Ostroff, editor of the *National Union Catalog of Manuscript Collections*, has been entering personal and corporate name authority records into the *Library of Congress Name Authorities* file for well over a year. This means that the foundation of historical names gradually will build within this critically important data base.

My suggestions about extending the concept of authority control were not original, but taken from papers and articles written by David Bearman, Richard Szary, Max Evans, and Richard Lytle. The bibliography entitled "Writings on Authority Control for Archival and Manuscript Materials" (at the end of this paper) demonstrates that archival literature devoted to the topic barely fills one page. You will also notice that if it were not for Bearman, most of the titles would not exist. Bearman alone, and with Szary and Lytle, writes about all the topics I briefly touched upon — the provenance approach, access points of form of material and function, multiple peer files, cultural data bases, and RLG's Seven States Project. Szary's paper compares the library bibliographic approach to the archival provenance approach. Max Evans analyzes the problems with the record group concept and concludes by explaining how agency history records (or authority control) helps solve these problems.

Suffice it to say, there has not been a great deal written yet about authority control in archives. I say "yet" because I think that the concept of authority control is one of the most provocative, innovative, and exciting ideas to emerge in a long time.

Which leads me to the final question: "Where does authority control fit in the context of the archival professional agenda?"

The Society of American Archivists (SAA) constituted a Task Force on Goals and Priorities (GAP) in 1982 to produce a systematic and comprehensive statement of what archivists want to accomplish as a profession. The task force (now an SAA Standing Committee entitled the Committee on Goals and Priorities, or CGAP) has the responsibility to turn this written planning document into focused, constructive, directed action. Although the report enumerated a large number of goals, objectives, and activities, the report also chose twelve priority areas.

One of the twelve priorities is to "develop and link manual and automated data bases about archival holdings on institutional, regional, national, and international levels." CGAP has sponsored a group of planning meetings in five crucial areas, and automation (both automated techniques and machine-readable records as an archival medium) is one of those five areas. As a result of the automation meeting, the development of archival descriptive standards and controlled vocabularies is on the CGAP agenda.

The Program Officer for Automation's position at SAA (my job), funded by the National Endowment for the Humanities, strongly supports archival information exchange and the development of standards, for part of the responsibility is to administer workshops that teach the MARC AMC format. New tasks include developing a workshop to teach library standards of the *Anglo-American Cataloging Rules*, 2d ed. (AARC2) and the *Library of Congress Subject Headings* (LCSH) to archivists and overseeing the revisions to Steven Hensen's *Archives, Personal Papers and Manuscripts.*

Perhaps most important are the activities that took place at SAA's 1987 annual meeting in New York City. The Description Section unanimously voted to ask the SAA Council to establish a funded task force to identify areas that need archival descriptive standards and to begin to develop such standards. Although this task force's mechanics have yet to be established, the development of standards, controlled vocabularies, and archival guidelines for use are on the minds of many.[6]

Though some have taken steps to put archival description mechanisms and intellectual access to materials towards the top of the archival profession's agenda, several major questions must be addressed at the outset or much of the work will be in vain.

The first and perhaps most critical issue surrounds users of archival materials. We think we know who our users are but these are only, in truth, unverified impressions. We have little analytical research based on scientific methods and models that ask (let alone answer) the question "Who uses archival materials?" Related unanswered questions include: "How do researchers approach archival materials?' "How do researchers ask questions and find appropriate materials?" "Where does the archivist fit in?" "Does an archivist always need to act as mediator between researcher and materials?" How can we expect to build successful data bases of descriptions of archival and manuscript materials if we have no idea for whom we are building them? If we have no idea how people find materials? If we have no idea what the purpose of the archival catalog is? Our profession desperately needs analytical studies that examine how people gain access to archival materials.[7]

On a more concrete level, we need to identify whom, within the archival profession, has the responsibility to develop, promulgate, and maintain descriptive standards, controlled vocabularies, and authority files. Many people assumed that the SAA's Committee on Archival Information Exchange would assume this role, but that assumption is currently in question.[8]

At a minimum, the profession as a whole has yet to take the first step to identify which categories of information need to be under authority control. Obviously, each institution faces this question when implementing MARC AMC. If the institution is participating in a national bibliographic network such as RLG or OCLC, then certain standards are required. But the archival profession has never agreed on categories on a national level.

We need to evaluate our profession vis-a-vis the library and information communities. How do we fit in? Where do we fit in? What is the archival profession's relationship with the bibliographic utilities? With the Library of Congress? With vendors? If Bearman and Szary are correct in suggesting that there are concrete rewards and financial benefits for archivists in derivative authority work, then when do we work together and when do we go it alone?

I do not implore, nor would I want, archival work to cease until we answer these questions. But I do think it is critical to devise and fund activities, projects, surveys, and studies that address these issues and begin to formulate answers.

I'd like to close with the following thought. One of National Information Systems Task Force's (NISTF) unsung contributions to the profession, through the development of the data element dictionary and the MARC AMC format, is that it has helped us to distinguish categories of information previously lumped together. We now recognize that we collect and generate information about the physical aspects of materials, the intellectual content, actions (what we do to the materials), and authority data. The ability to separate these categories has opened up new options and possibilities, with a potential that appears powerful. The very fact that this seminar is taking place is an exciting event for the profession, and I, for one, look forward to seeing many positive results.

NOTES

1. Ghikas, Mary W., ed., *Authority control: the key to tomorrow's catalog. Proceedings of the 1979 Library and Information Technology Association Institutes.* Phoenix, Ariz.: Oryx Press, 1982.

2. Hensen, Steven L. "A modern theory of manuscript cataloging and the master record of manuscript collections." Paper delivered at the annual meeting of the Society of American Archivists, Berkeley, California, September 1, 1981, pp. 4-5. Although Hensen talks about manuscript collections, I think the analogy extends to archival materials as well.

3. See bibliography.

4. Bearman, David. "Building the foundations of national information systems for archives and manuscript repositories: the RLIN seven states project." Paper delivered at the annual conference of the National Association of Government Archives and Records Administrators (NAGARA), July, 1987, pp. 12.

5. Miller, Bruce R. "Authority control in the network environment." In *Authority control*, Ghikas, ed., pp. 38-9.

6. Subsequently, Lawrence Dowler, Harvard University, submitted an application to the National Historical Publications and Records Commission, to support a work group of archivists, representing various interests in the profession, to consider questions relating to descriptive standards and to provide a conceptual framework within which the profession can consider these issues. The Commission will act upon this proposal in June, 1988.

7. For a thoughtful and provoking discussion of these issues see: Dowler, Lawrence. "The availability and use of records: a research agenda." Paper delivered at the annual meeting of the Society of American Archivists, New York, New York, September, 1987.

8. See note 6.

WRITINGS ON AUTHORITY CONTROL FOR ARCHIVAL AND MANUSCRIPT MATERIALS

Bearman, David. "'Who About What?', or 'From Whence, Why & How:' Establishing Intellectual Control Standards to Provide for Access to Archival Materials." In *Archives, Automation and Access*, 39-47. Vancouver: University of British Columbia, 1986.

_____. "Buildings As Structures, As Art, and As Dwellings: Data Exchange Issues in an Archival Information Network." Proceedings of the International Conference on Data Bases in the Humanities and Social Sciences, Auburn University, Montgomery, Alabama, July 11-13, 1987s (forthcoming in spring, 1988).

_____. "Working Papers for the 1986 Research Fellowship Program for Study of Modern Archives," administered by the Bentley Historical Library, University of Michigan, (forthcoming in summer, 1988).

_____. "Building the Foundations of National Information Systems for Archives and Manuscript Repositories: The RLIN Seven States Project." Paper presented at the annual conference of the National Association of Government Archives and Records Administrators (NAGARA), July, 1987.

Bearman, David, and Richard Lytle. "The Power of the Principle of Provenance," *Archivaria* 21 (Winter 1985-86): 14-27.

Bearman, David, and Richard Szary. "Beyond Authorized Headings: Authorities as Reference Files in a Multi-Disciplinary Setting." In *Authority Control Symposium: Occasional Papers No. 6*. Papers presented during the 14th Annual ARLIS/NA Conference, New York, N.Y. February 10, 1986, edited by Karen Muller. Tuscon: Art Libraries of North America, 1987, pp. 69-78.

Evans, Max. "Authority Control: An Alternative to the Record Group Concept," *American Archivist* 49 (Summer 1986): 249-261.

Szary, Richard. "Expanding the Role of Authority Files in the Archival Context." Paper presented at the Society of American Archivists annual meeting, November, 1985.

AUDIENCE DISCUSSION

PETER HIRTLE:

I hate to ask such a nuts-and-bolts question when you leave us with so many big issues to discuss. I heard you call for authority data bases that would include a much broader range of information than we currently include. And I agree that would be nice. But, I just heard Marion Matters say that at the Minnesota Historical Society they can't even establish authority control over the names they are using in their catalog records. They establish authority control only for the Minnesota ones. So, what are we going to cut? Are we going to cut our acquisitions in half and our processing in half in order to do authority work and make authority files?

LISA WEBER:

I think one of the motivating factors behind the ideas that David Bearman, Rich Szary, and Richard Lytle talk about is that if we could be creating authority files that are actually reference files, there would be some economic benefit in that other people would want to use them. At lunch, Harry Heiss asked me if anybody could make a profit doing this? Would vendors be interested in these files? In adjacent communities, vendors have shown interest. The Library of Congress provides tremendous resources for the library community, and the Getty has funded the development of the Art and Architecture Thesaurus Project.

DAVID BEARMAN:

We had another discussion at lunch, about the library community's bibliographic data bases. It's my position that the library community's bibliographic data bases are actually authority files because there are lots of copies of the book out there, but there is only one authority record. And, in fact, they have been doing quite a good job of not only building these authority files, but building whole networks to exchange these authority records, and there are a lot of people making money selling these authority records, both in the networks and now in all kinds of new packages like CD Roms. And if you look at those records, you discover that the specific expertise of the library community is to develop reference files about bibliographic entities, and the specific expertise of the archival community is to develop reference files about the world of corporate entities. Archivists, however, break up all the information they have and bury it in unmanageable files called finding aids, each one of which contains one record and is inaccessible. And, in fact, what Marion Matters points to, and I think what comes out of the discussions in the archival community, is that we have, as a profession, a significant advantage in creating ourselves shared corporate reference files similar to the shared bibliographic reference files that the library community constructs. I think Jackie Dooley stated this most clearly, in saying that only the people who are there with those records can really build the information about the local corporate entities. Just as there is shared benefit in archivists constructing corporate entity data bases, there are other communities that create primary reference files of geographic entities, geopolitical entities, and forms, (by which I mean physical material types). One of the potentials when you consider reference files constructed by various communities, is that each of these communities has a product that it sells to another community. So, there is some potential here, and I think that is one of the reasons why within the Smithsonian there is the potential for the library to benefit from the work of the archival units, the archival units to benefit from the work of the curatorial units, and the curatorial units to benefit from each other's work.